VECTOR-VALUED INVERSE PROGRAMMING

by

DAVID F. GRAY

A thesis submitted in partial fulfilment of the requirement for the Degree of Doctor of Philosophy at Dalhousie University, Halifax, Nova Scotia, April, 1977.

Table of Contents

Abstract

In this thesis we consider pairs of mathematical programming problems with optimal values that are functionally inversely related. We study the application of inverse programming to vector-valued mathematical programming problems. This leads to a consideration of the relationship between inverses and duals and also to methods of obtaining solutions for vector-valued problems that are efficient with respect to the constraints as well as to the objectives.

Acknowledgements

The author wishes to express his thanks to his supervisor, Dr. W.R.S. Sutherland, who not only introduced him to the problem on which the thesis is based but gave him much help and encouragement during his years as a graduate student. The author also thanks Dr. J. Borwein, Dr. W. Smith and Dr. F. Fiala for reading the thesis and giving helpful advice. Finally he wishes to thank his wife Linda for typing this ~~thesis~~.

Introduction

In this thesis we consider pairs of mathematical programming problems with optimal values that are functionally inversely related. The idea of inverse programming was first introduced in [5] to provide a solution method for a problem that was difficult to solve by other means. We outline this example but do not pursue this particular use of inverse programming any further. Instead we consider its application to vector-valued mathematical programming problems. This leads to a consideration of the relationship between inverses and duals and also to methods of obtaining solutions for vector-valued problems that are efficient with respect to the constraints as well as to the objectives.

In chapter 1 we develop a definition of an inverse pair of mathematical programming problems. We also introduce the notation and symbols used in the rest of the thesis. In chapter 2 we review some general results on solutions of vector-valued programming problems and consider the restriction of these to the linear case. In chapter 3 we discuss linear inverse problems; first those in which one problem is scalar-valued and one is vector-valued and then those in which both problems are vector-valued. In chapter 4 we review some duality results in nonlinear programming and apply these to the problem of finding solutions to nonlinear inverse pairs. In doing this we develop a saddle function that treats the objective functions and constraints more symmetrically than the saddle function usually used.

The most important results in this thesis are theorems 4-14, 4-15 and 4-16 which summarize the relationship between inverse problems and their duals in nonlinear programming. These allow us to develop methods of obtaining solutions to vector-valued problems which are efficient with respect to both objectives and constraints. For the linear case we develop algorithms (3-6, 3-7, 3-13) for these methods based on the restricted versions of these results given in theorems 3-5 and 3-11 and results particular to the linear case given in theorems 3-4 and 3-12. Before theorem 4-14 can be proved it is necessary to develop a framework for the problem. Definition 3-1 provides the first rigorous definition of an inverse pair of vector-valued problems. Definition 4-13 introduces a symmetric Lagrange function that is particularly useful when considering inverse pairs and may be useful in other areas of mathematical programming as well.

Throughout the thesis we indicate interesting areas for further research. In particular we should be able to provide algorithms for other specialized classes of problems similar to those developed for the linear case. Also we should consider other uses of the symmetric Lagrange function.

We number theorems, lemmas, definitions, and algorithms sequentially within each chapter. Each is coded with the chapter and sequence number. Hence theorem 2-3 would be the third numbered item in chapter 2. The end of any of the above is marked by a ▮ as is the commencement of a proof. References to the bibliography are given in square brackets. Hence [13] refers to the thirteenth entry in the bibliography.

Chapter One - Inverse Problems: Definition and Background

In this chapter we develop the background definitions and the notation needed for the results of this thesis. The main definition is 1-3 which rigorously defines an inverse pair of mathematical programming problems. We then discuss two examples which show the usefulness of this concept and clarify the definition. Finally we define several restricted inverse pairs of problems that are considered in the following chapters.

Since we discuss vector-valued problems we must consider orders on elements in R^n, the n-dimensional space of vectors with real components. In R^1 we distinguish two inequalities: $x > y$ (x is strictly greater than y) and $x \geqq y$ (x is greater than or equal to y). In the more general R^n we have three possible inequalities. These we denote as follows:

$$\text{for } x = (x_1,\ldots,x_n),\ y = (y_1,\ldots,y_n)$$

$$\begin{aligned}
\text{then } x \geqq y &\text{ if } x_i \geqq y_i \text{ for } i = 1,2,\ldots,n; \\
x \geq y &\text{ if } x_i \geqq y_i \text{ for } i = 1,2,\ldots,n \\
&\text{ and } x_j > y_j \text{ for some } j; \\
x > y &\text{ if } x_i > y_i \text{ for } i = 1,2,\ldots,n;
\end{aligned}$$

to be read as:

$\geqq$ "greater than or equal to",

$\geq$ "greater than but not equal to",

$>$ "strictly greater than".

We are interested in the maximum or minimum element, if it exists, in some subset $S \subset R^n$. If $n = 1$, the common notion of a maximum (or minimum) suffices, that is:

$$x_0 = \max\{x \in S\} \text{ if and only if}$$
$$x_0 \in S \text{ and } x_0 \geqq x \text{ for all } x \in S.$$

If $n > 1$, we may not be able to compare all vectors in S so this notion is not sufficient for our needs. Instead we seek a Pareto maximum (or minimum) defined as follows.

<u>Definition 1-1</u>: $x = (x_1, \ldots, x_n)$ S is a <u>Pareto</u> <u>maximum</u> if, for any $y \in S$,

$$y \geqq x \text{ implies } y = x.$$

We write this as

$$x = \text{Pmax}\{y\}$$
$$\text{such that } y \in S.$$
$$\text{or } x = \text{Pmax}\{y \in S\}.$$

We define a <u>Pareto</u> <u>minimum</u> in a similar fashion and denote it by "Pmin". A Pareto maximum or minimum of S is also called an <u>efficient point</u> of S or a <u>Pareto</u> <u>optimum</u>. ∎

Notice that if $n = 1$ a Pareto optimum is optimal in the usual sense, so in general we can look for Pareto optima for any n.

The next definition defines the basic set that we are interested in, two problems on that set and some useful subsets.

<u>Definition 1-2</u>: (1) $T \subset R^m x R^n$ is a set of ordered pairs

(a,b) such that $a \in R^m$ and $b \in R^n$ with the following properties:

$$\{(\bar{a},\bar{b}) \text{ and } (\hat{a},\hat{b}) \in T \text{ and } \hat{b} \geq \bar{b}\} \Rightarrow \{(\bar{a},\hat{b}) \in T\}$$

$$\{(\bar{a},\bar{b}) \text{ and } (\hat{a},\hat{b}) \in T \text{ and } \hat{a} \leq \bar{a}\} \Rightarrow \{(\hat{a},\bar{b}) \in T\}$$

(2) We define two problems:

$$P(\bar{b}) \quad \text{find } \bar{a} = \text{Pmax}\{a\} \quad \text{such that } (a,b) \in T,\ b \leq \bar{b},$$

and

$$I(\bar{a}) \quad \text{find } \bar{b} = \text{Pmin}\{b\} \quad \text{such that } (a,b) \in T,\ a \geq \bar{a}.$$

(3) We define two general sets A and B:

$A = \{a \mid (a,b) \in T \text{ for some } b \in R^n\}$,

$B = \{b \mid (a,b) \in T \text{ for some } a \in R^m\}$.

(4) We define two subsets of A and B:

$\bar{A} = \{a \in A \mid \text{ for some } b \in B \nexists\ \bar{a} \in A, \bar{a} > a, (a,b) \in T\}$,

$\bar{B} = \{b \in B \mid \text{ for some } a \in A \nexists\ \bar{b} \in B, \bar{b} < b, (a,b) \in T\}$.

(5) We define sets $A(\bar{b})$ and $B(\bar{a})$ depending on $\bar{b}$ and $\bar{a}$ as:

$A(\bar{b}) = \{\bar{a} \mid \bar{a}$ is a Pareto optimum in $P(\bar{b})\}$,

$B(\bar{a}) = \{\bar{b} \mid \bar{b}$ is a Pareto optimum in $I(\bar{a})\}$. ∎

We next restrict the set T so that $P(\bar{b})$ and $I(\bar{a})$ have optimal values that are inversely related. This definition is the basis for all our results.

<u>Definition 1-3</u>: We call $P(\bar{b})$ and $I(\bar{a})$ an <u>inverse pair</u> of mathematical programming problems if T is a

nonempty closed set, if $A(\bar{b})$ and $B(\bar{a})$ contain finite elements for all $\bar{a} \in A$ and $\bar{b} \in B$, and if the following monotonicity property holds:

(1) if $\bar{b}$ and $\hat{b} \in \bar{B}$ then

$$\{\hat{b} > \bar{b} \text{ and } \bar{a} \in A(\bar{b})\} \Rightarrow \{\hat{a} \in A(\hat{b}) \mid \hat{a} > \bar{a}\},$$

(2) if $\bar{a}$ and $\hat{a}$ $\bar{A}$ then

$$\{\hat{a} < \bar{a} \text{ and } \bar{b} \in B(\bar{a})\} \Rightarrow \{\hat{b} \in B(\hat{a}) \mid \hat{b} < \bar{b}\}. \blacksquare$$

The meaning of this definition is clarified in the discussion below but notice that the conditions state that if we relax all the restrictions in either problem then we can find a new efficient point which shows an improvement in every coordinate.

If $P(\bar{b})$ and $I(\bar{a})$ form an inverse pair we distinguish between them by calling $P(\bar{b})$ the primal problem and $I(\bar{a})$ the inverse problem. However this distinction is somewhat artificial. Consider the following lemma.

<u>Lemma 1-4</u>: Let $a' = -a$, $b' = -b$. Define

$$T' = \{(b',a') \text{ such that } (a,b) \in T\},$$

$$A' = \{a' \mid (b',a') \in T' \text{ for some } b' \in R^n\},$$

$$B' = \{b' \mid (b',a') \in T' \text{ for some } a' \in R^m\},$$

$$\bar{A}' = \{a' \in A' \mid \text{for some } b' \in B' \not\exists\ \bar{a}' \in A', \bar{a}' < a', (b',a') \in T'\},$$

$$\bar{B}' = \{b' \in B' \mid \text{for some } a' \in A' \not\exists\ \bar{b}' \in B', \bar{b}' > b', (b',a') \in T'\}.$$

Then if $P(\bar{b})$ and $I(\bar{a})$ are inverse problems so are

$P'(\bar{a}')$ find $\bar{b}' = \text{Pmax}\{b'\}$

such that $(b',a') \in T'$

$a' \leq \bar{a}'$,

and $I'(\bar{b}')$ find $\bar{a}' = \text{Pmin}\{a'\}$

such that $(b', a') \in T'$

$b' \geq \bar{b}'$.

Proof:

Since T is a nonempty closed set so is T'. $A' = -A$, $\bar{B}' = -B$, $\bar{A}' = -\bar{A}$ and $\bar{B}' = -\bar{B}$. $P'(\bar{a}')$ is equivalent to $I(\bar{a})$ and $I'(\bar{b}')$ is equivalent to $P(\bar{b})$. Hence the existence of finite elements of $A(\bar{b})$ and $B(\bar{a})$ implies the existence of finite elements of $A'(\bar{b}')$ and $B'(\bar{a}')$, the efficient sets for the new problems. Now assume $\hat{a}', \bar{a}' \in \bar{A}'$ and $\hat{a}' > \bar{a}'$; since $a' = -a$ we must have $\hat{a}, \bar{a} \in \bar{A}$ and $\hat{a} < \bar{a}$. But $P(\bar{b})$ and $I(\bar{a})$ are an inverse pair so

$$\{\hat{a} < \bar{a} \text{ and } \bar{b} \in B(\bar{a})\} \Rightarrow \{\exists\, \hat{b} \in B(\hat{a}),\ \hat{b} < \bar{b}\};$$

or, transforming the a's and b's and noting that

$$B'(\bar{a}') = \{-\bar{b} \mid \bar{b} \in B(\bar{a})\},$$

then $\{\hat{a}' > \bar{a}' \text{ and } \bar{b}' \in B'(\bar{a}')\} \Rightarrow \{\exists\, \hat{b}' \in B'(\hat{a}') \mid \hat{b}' > \bar{b}'\}$. Similarly if $\hat{b}', \bar{b}' \in \bar{B}'$, then

$$\{\hat{b}' < \bar{b}' \text{ and } \bar{a}' \in A'(\bar{b}')\} \Rightarrow \{\exists\, \hat{a}' \in A'(\hat{b}') \mid \hat{a}' < \bar{a}'\}.$$

Thus $P'(\bar{a}')$ and $I'(\bar{b}')$ form an inverse pair of problems with $P'(\bar{a}')$ the primal and $I'(\bar{b}')$ the inverse.■

Thus it is actually immaterial which problem we consider to be primal.

We also use one other equivalent pair of inverse problems. If we replace $-b$ by b' then the original pair of problems is equivalent to:

$$\begin{aligned} &\text{find } \bar{a} = \text{Pmax}\{a\} \\ &\quad \text{such that } (a,-b') \in \mathcal{T} \\ &\qquad\qquad\qquad\qquad b' \geq \bar{b}', \end{aligned}$$

$$\begin{aligned} \text{and} \quad &\text{find } \bar{b}' = \text{Pmax}\{b'\} \\ &\quad \text{such that } (a,-b') \in \mathcal{T} \\ &\qquad\qquad\qquad\qquad a \geq \bar{a}. \end{aligned}$$

In this pair the symmetry is even more transparent. Which pair we use does not matter because of the above equivalence but one may be more natural in a particular example or yield a less confusing proof.

We now consider the meaning and intent of definition 1-3. Let f be a function from the real numbers into the real numbers, $f:R\to R$. Then, if f is continuous and strictly monotone, f has a unique inverse function $g:R\to R$ such that $f(b) = a \iff g(a) = b$ (see figure 1-1). If we let $\mathcal{T} = \{(a,b) \mid a \leq f(b)\}$ or, equivalently, $\mathcal{T} = \{(a,b) \mid b \geq g(a)\}$ then the optimal solution to $P(\bar{b})$ is $\bar{a} = f(\bar{b})$, i.e. $A(\bar{b}) = \{f(\bar{b})\}$, and the optimal solution to $I(\bar{a})$ is $\bar{b} = g(\bar{a})$, i.e. $B(\bar{a}) = \{g(\bar{a})\}$. It is easy to see that $P(\bar{b})$ and $I(\bar{a})$ are an inverse pair of problems. For this pair the important part of $\mathcal{T}$ is the curve $a = f(b)$. This leads to the following definition.

<u>Definition 1-5</u>: The <u>active boundary</u> of $\mathcal{T}$, called $\mathcal{R}$ is defined as

$$\mathcal{R} = \{(a,b) \in \mathcal{T} \mid a \in A(b) \text{ or } b \in B(a)\}.\blacksquare$$

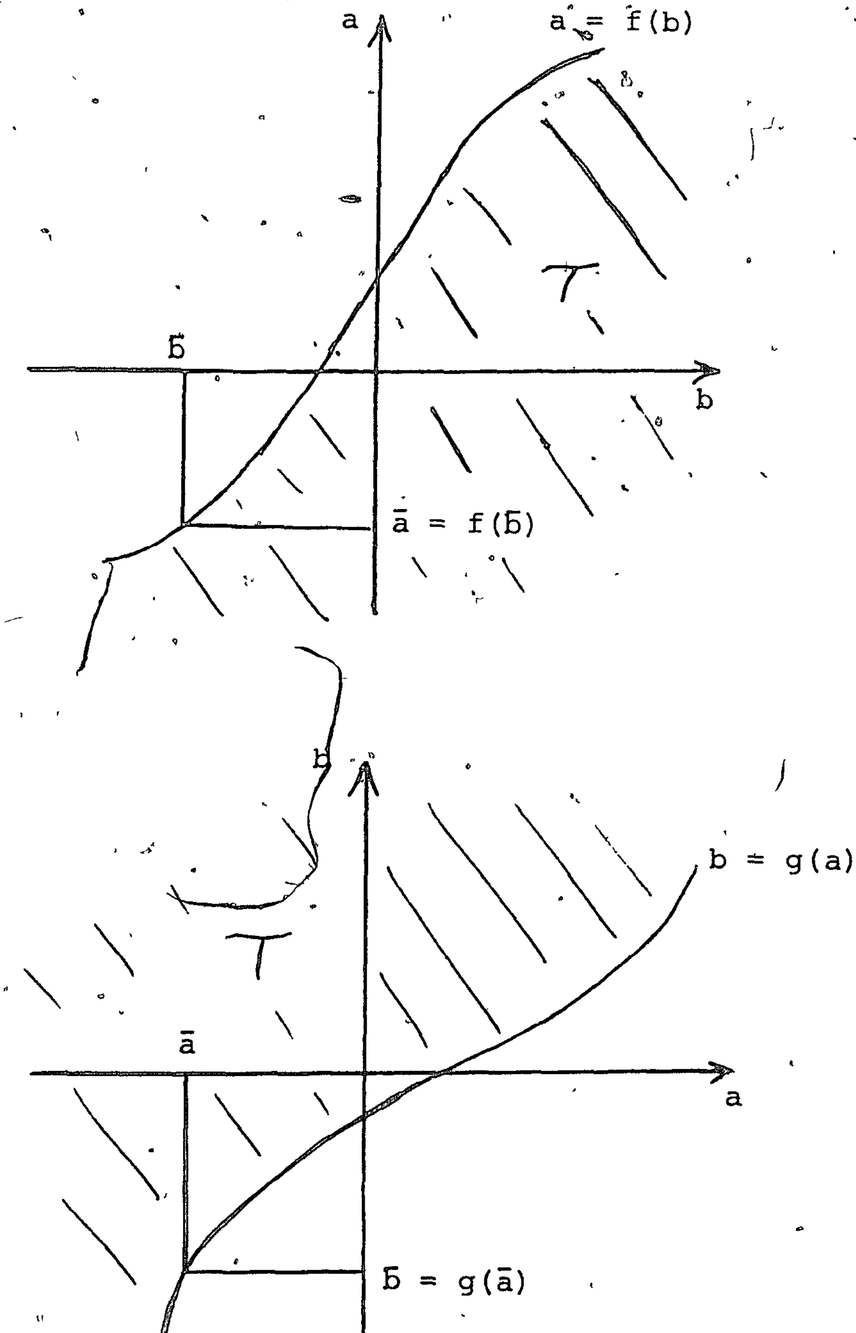


Figure 1-1: A pair of scalar inverse functions.

By this definition, definition 1-3 implies that the active boundary for a problem in which a and b are scalar-valued is strictly monotone.

There are several ways to extend these ideas to mappings from R^m to R^n. Definition 1-3 is an attempt to construct the weakest generalization that yields useful results for the problems we consider.

One way to generalize the inverse relationship depicted in figure 1-1 is through the property of T having an active boundary that is strictly increasing in each component when considered as giving a as a function of b or b as a function of a. However in this case we get the following result.

<u>Lemma 1-6</u>: If $P(\bar{b})$ and $I(\bar{a})$ are an inverse pair of problems for $T \subset R^m x R^n$ and if T has a strictly monotone (as described above) active boundary, then

(1) (i) if $\hat{b}, \bar{b} \in B$ then
$\{\hat{b} \geq \bar{b}$ and $\bar{a} \in A(\bar{b})\} \Rightarrow \{\exists\ \hat{a} \in A(\hat{b}) \mid \hat{a} \geq \bar{a}\}$,

(ii) if $\hat{a}, \bar{a} \in A$ then
$\{\hat{a} \leq \bar{a}$ and $\bar{b} \in B(\bar{a})\} \Rightarrow \{\exists\ \hat{b} \in B(\hat{a}) \mid \hat{b} \leq \bar{b}\}$,

(2) (i) if $\hat{b} \in B(\bar{a})$ then $\bar{a} \in A(\hat{b})$

(ii) if $\hat{a} \in A(\bar{b})$ then $\bar{b} \in B(\hat{a})$.

<u>Proof</u>:

(1) (i) $(\bar{a},\bar{b}) \in T$. Therefore $\bar{a}$ is feasible in $P(\hat{b})$ since $\hat{b} \geq \bar{b}$. If $\bar{a}$ is efficient in $P(\hat{b})$ then $(\bar{a},\hat{b})$ is on the active boundary. But $\hat{b} \neq \bar{b}$, so if $(\bar{a},\bar{b})$ and $(\bar{a},\hat{b})$

are on the active boundary it is not strictly monotone. Therefore $\exists\ \hat{a} \in A(\hat{b})$ such that $\hat{a} \geq \bar{a}$ and implication (1) (i) is proved.

(1) (ii) is proved similarly.

(2) (i) Let $\hat{b} \in B(\bar{a})$ and assume $\bar{a} \notin A(\hat{b})$. Then $\exists\ \tilde{a} \in A(\hat{b})$ such that $\tilde{a} \geq \bar{a}$. Therefore $(\bar{a},\hat{b})$ and $(\tilde{a},\hat{b})$ are elements of R. But, as above, this contradicts the fact that R is strictly monotone. Therefore $\bar{a} \in A(\hat{b})$.

(2) (ii) is proved similarly.■

These properties would put strong restrictions on the problems that could be considered as we point out in the examples below. Thus a weaker notion of monotonicity is needed.

Definition 1-3 provides the needed weaker monotonicity. This monotonicity requirement is weaker in two ways. First, in definition 1-3 we only require the condition to hold for b's in $\bar{B}$ and a's in $\bar{A}$. This is a minor adjustment and we normally are interested only in values in these sets. However it does allow us to treat a large class of linear programming problems without any difficulty as is shown in the examples later in the chapter. Second, the stronger monotonicity of lemma 1-6 guarantees an a with at least one larger component if any component of b is increased while the monotonicity defined in definition 1-3, though it does guarantee an a with all components larger, requires increases in all components of b. To see that this actually is weaker we must consider what we can say if $\hat{b}$, and $\bar{b}$

are in $\mathcal{B}$ and we only have the conditions of definition 1-3.

<u>Lemma 1-7</u>: If $\mathcal{T}$, $P(\bar{b})$, $I(\bar{a})$ satisfy definition 1-3 then

(1) (i) if $\hat{b}, \bar{b} \in \mathcal{B}$ then

$\{\hat{b} \geq \bar{b}$ and $\bar{a} \in A(\bar{b})\} \Rightarrow \{\exists\ \hat{a} \in A(\hat{b})\ |\ \hat{a} \geqq \bar{a}\}$

(ii) if $\hat{a}, \bar{a} \in A$ then

$\{\hat{a} \leq \bar{a}$ and $\bar{b} \in B(\bar{a})\} \Rightarrow \{\exists\ \hat{b} \in B(\hat{a})\ |\ \hat{b} \leqq \bar{b}\}$,

(2) (i) if $\hat{b} \in B(\bar{a})$ then $\exists\ \hat{a} \in A(\hat{b})$, $\hat{a} \geqq \bar{a}$

(ii) if $\hat{a} \in A(\bar{b})$ then $\exists\ \hat{b} \in B(\hat{a})$, $\hat{b} \leqq \bar{b}$.

<u>Proof:</u>

▮ (1) (i) $\bar{a} \in A(\bar{b}) \Rightarrow \exists\ (\bar{a},\tilde{b}) \in \mathcal{T}$ with $\tilde{b} \leqq \bar{b}$, but $\bar{b} \leq \hat{b}$ so $\tilde{b} \leq \hat{b}$ and $\bar{a}$ is a feasible solution to $P(\hat{b})$. Thus, if $\bar{a}$ is not efficient, there is an $\tilde{a} \geq \bar{a}$ that is. In either case $\exists\ \hat{a} \in A(\hat{b})$, $\hat{a} \geqq \bar{a}$.

(1) (ii) is proved similarly.

(2) (i) If $\hat{b} \in B(\bar{a})$ then $\exists\ \tilde{a}$ such that $(\tilde{a},\hat{b}) \in \mathcal{T}$ and $\tilde{a} \geqq \bar{a}$. Since $(\tilde{a},\hat{b}) \in \mathcal{T}$ then $\tilde{a}$ is feasible in $P(\hat{b})$, so there is an $\hat{a} \in A(\hat{b})$, $\hat{a} \geqq \tilde{a} \geqq \bar{a}$.

(2) (ii) is proved similarly.▮

Thus if we have an inverse pair of problems but do not have a strictly monotone active boundary, increasing one component of b may not give us any increase in the efficient solutions to the primal problem. Though in these two respects the monotonicity of lemma 1-6 is stronger, it may not imply the monotonicity of definition 1-3. For example, let $b \in R^2$ and $a \in R^2$ and consider the primal problem. If we increase either component of b (and stay in $\mathcal{B}$) we can obtain an efficient a that may only show an improvement in one component. If increases to both components of b, one at a time affect the same component of a, then increasing both at the same time may still only produce an increase in one component of a. In this case the problems would not be an inverse pair. Thus the monotonicity of lemma 1-6 and the monotonicity of definition 1-3 cannot be put in a hierarchy though they both imply the monotonicity of lemma 1-7.

We now consider two examples that show the usefulness of our definition of an inverse pair of problems.

<u>Example 1-8</u>: Consider the standard linear programming example of a firm choosing levels of production of various processes so as to maximize its profit from the use of a given amount of resources. This is an "Activity Analysis of Production" model as given in Koopmans [19]. That is:

let a be total profit,

b be the given resource vector,

c be the unit profit vector,

D be the technology matrix,

then the standard LP problem is:

$$\begin{aligned} &\text{maximize} \quad a \\ &\text{subject to} \quad cx \geqq a \\ &\qquad\qquad\quad Dx \leqq \bar{b} \\ &\qquad\qquad\quad\ \ x \geqq 0. \end{aligned}$$

Now we define $T = \{(a,b) \mid \exists\, x \geqq 0,\ cx \geqq a,\ Dx \leqq b\}$ so the above problem is equivalent to:

$$P(\bar{b}) \qquad \begin{aligned} &\text{maximize} \quad a \quad (\text{or find } \bar{a} = \text{Pmax}\{a\}) \\ &\quad\text{such that} \quad (a,b) \in T \\ &\qquad\qquad\qquad\quad b \leqq \bar{b}. \end{aligned}$$

Since b is a vector, the inverse problem on T is:

$$I(\bar{a}) \qquad \begin{aligned} &\text{find} \quad \bar{b} = \text{Pmin}\{b\} \\ &\quad\text{such that} \quad (a,b) \in T \\ &\qquad\qquad\qquad\quad a \geqq \bar{a}. \end{aligned}$$

Here we are trying to find a minimal vector of quantities of resources needed to produce a desired output.

If c and D are positive $\bar{b}$ must be nonnegative if $P(\bar{b})$ has any feasible solutions. If $\bar{b} \geq 0$, $\bar{x}$ solves $P(\bar{b})$ and $\hat{b} > \bar{b}$, there is an $\alpha > 1$ such that $\alpha\bar{b} < \hat{b}$ and $\alpha\bar{x}$ solves $P(\alpha\bar{b})$. Thus $\alpha\bar{x}$ is feasible in $P(\hat{b})$ and, if $\hat{a}$ is the value of $P(\hat{b})$, $\hat{a} \geqq C\alpha\bar{x} = \alpha C\bar{x} > C\bar{x}$. If $\bar{b} = 0$, the optimal value of $P(\bar{b})$ is zero. If $\hat{b} > \bar{b}$, $\hat{b} > 0$ and the optimal value of $P(\hat{b}) > 0$. We can argue similarly for the inverse problem if $\bar{a} \geqq 0$. In this example $B = \bar{B} = \{b \geqq 0\}$ while $A = R$ and $\bar{A} = R^{+}$. Thus the conditions of definition 1-3 are satisfied for the pair of problems $P(\bar{b})$ and $I(\bar{a})$ if c and D are positive. A and $\bar{A}$ differ in this case. Though the only really interesting problems have $a \geq 0$, $I(a)$ is solvable for $a < 0$. As definition 1-3 stands, we can handle this example without adding extra restrictions.

If definition 1-3 holds we now have an inverse pair of problems with the following properties. If we increase the supply of all resources from $\bar{b}$ to $\hat{b}$, $\hat{b} > \bar{b}$, then $P(\hat{b})$ has a solution that is more profitable than the optimal solution to $P(\bar{b})$ - a natural enough restriction. However if we increase the supply of some but not all of the resources, so $\hat{b} \geq \bar{b}$, then, as lemma 1-7 shows, the new optimal solution is as profitable as the old, but the profit may not actually increase. Thus our model allows for excesses to occur in the supply of some resources. Lemma 1-6 shows that this could not occur if we use the more restrictive form of monotonicity of a mapping between R^m and R^n. This is one reason for choosing the weaker version for definition 1-3.

For the inverse problem, if we decrease the amount of profit that is desired, the conditions of definition 1-3 imply that there is a new efficient solution that uses less of each resource - again a natural enough restriction. This second problem is not a standard LP problem since it has a vector-valued objective. The existence of solutions to problems of this type and methods of finding them are discussed in the next chapter. The particular example mentioned here is a problem of increasing importance in economics because of the recent prominence of ideas of controlled growth and resource conservation.∎

<u>Example 1-9</u>: This example from Cassidy [4] initially lead to the consideration of inverse pairs of optimization problems in Cassidy, Field and Sutherland [5]. This is a problem in random payoff games. We assume that the players face a payoff matrix $A = \{a_{ij}\}$ with a_{ij}'s that are random with known distribution function $P(a_{ij} \geq \beta)$.

The row player may view his problem from two different perspectives. First, he can consider maximizing the return he gets (β) with at least a given probability $(\bar{\alpha})$, i.e. his problem is:

$$\begin{aligned} &\text{maximize} \quad \beta \\ &\quad \text{such that} \quad \sum_{i=1}^{n} x_i P(a_{ij} \geq \beta) \geq \bar{\alpha} \quad \forall\, j \\ &\qquad\qquad\qquad\qquad \sum_{i=1}^{n} x_i = 1 \\ &\qquad\qquad\qquad\qquad x_i \geq 0 \quad \forall\, i. \end{aligned}$$

Second, he can consider maximizing the probability (α) of getting at least a given return $(\bar{\beta})$, i.e. his problem is:

$$\begin{aligned} &\text{maximize} \quad \alpha \\ &\quad \text{such that} \quad \sum_{i=1}^{n} x_i P(a_{ij} \geq \bar{\beta}) \geq \alpha \quad \forall\, j \\ &\qquad\qquad\qquad\qquad \sum_{i=1}^{n} x_i = 1 \\ &\qquad\qquad\qquad\qquad x_i \geq 0 \quad \forall\, i. \end{aligned}$$

Now we define $T = \{(\beta,\alpha) \mid \exists\, x \geq 0,\ \sum_{i=1}^{n} x_i = 1,\ \sum_{i=1}^{n} x_i P(a_{ij} \geq \beta) \geq \alpha \ \ \forall\, j\}$. Then the above problems become, respectively:

$$P(\bar{\alpha}) \qquad \begin{array}{ll} \text{maximize} & \beta \\ \text{such that} & (\beta,\alpha) \in T \\ & \alpha \geq \bar{\alpha}, \end{array}$$

$$I(\bar{\beta}) \qquad \begin{array}{ll} \text{maximize} & \alpha \\ \text{such that} & (\beta,\alpha) \in T \\ & \beta \geq \bar{\beta}. \end{array}$$

In [5], the authors show that if the distributions of the a_{ij} are strictly monotone, then definition 1-3 is satisfied and $P(\bar{\alpha})$ and $I(\bar{\beta})$ are an inverse pair. Notice though that $I(\bar{\beta})$ is a linear programming problem while $P(\bar{\alpha})$ is extremely nonlinear since the x_i and β are all variables. Here the inverse relationship allows us to solve $P(\bar{\alpha})$ by repeated solution of $I(\bar{\beta})$ as discussed in the cited paper.■

Notice that in both example 1-8 and example 1-9 one of the problems is a linear programming problem. This shows that the inverse of a mathematical programming problem is not unique.

We now consider some special cases of T that are useful in the following chapters. The easiest situation to handle is one in which both $P(\bar{b})$ and $I(\bar{a})$ are scalar-valued as in example 1-9. If this is the case we say that we have a scalar inverse pair of mathematical programming problems. In example 1-8, both the objectives and constraints of the problems are linear. We can generalize example 1-8, keeping linearity but allowing a to be vector-valued. Now T takes the form

$T = \{(a,b) \mid \exists\ x \in K,\ Dx \leqq b,\ Cx \geqq a,\ K$ some convex set, D an $m\times n$ matrix, C a $k\times n$ matrix$\}$. K is usually the non-negative orthant. In this case we say we have a <u>linear inverse pair</u> of programming problems. If we again generalize the above definition of T somewhat so

$T = \{(a,b) \mid \exists\ x \in K,\ g(x) \leqq b,\ f(x) \geqq a,\ g:R^n \to R^m,\ f:R^n \to R^k\}$

then $P(\bar{b})$ and $I(\bar{a})$ are both examples of standard (though possibly vector-valued) nonlinear programming problems. In this case we say we have a <u>nonlinear inverse pair</u> of programming problems. Note that this last definition is a restriction to definition 1-3 since we require that a and b can be specified by separate sets of constraints. Definition 1-3 allows cases where such a separation is not possible (as in example 1-9).

For linear and nonlinear inverse pairs where the underlying vector x is of importance we make the following definitions.

<u>Definition 1-10</u>: (a) If $\bar{a} \in A(\bar{b})$ and $\bar{b} \in B(\bar{a})$ such that $\bar{a} = C\bar{x}$ and $\bar{b} = D\bar{x}$ for $\bar{x} \in K$ in the linear case, or $\bar{a} = f(\bar{x})$ and $\bar{b} = g(\bar{x})$ for $\bar{x} \in K$ in the nonlinear case, then $(\bar{a},\bar{b},\bar{x})$ is called an <u>optimal inverse triple</u>.

(b) $\hat{a} \in A(\bar{b})$ and $\hat{b} \in B(\bar{a})$ are called <u>efficient values</u> for $P(\bar{b})$ and $I(\bar{a})$ respectively. If $\hat{x}$ is the solution that yields $\hat{a}$ (or $\hat{b}$), then $\hat{x}$ is called an <u>efficient solution</u> for $P(\hat{b})$ (or $I(\hat{a})$).

(c) $A(\bar{b})$ and $B(\bar{a})$ are called the set of <u>efficient values</u> for $P(\bar{b})$ and $I(\bar{a})$ respectively. The set of all $\bar{x}$ that yield the set of efficient values is called the <u>set of efficient solutions</u>.

(d) $A^I(\bar{b}) \subset A(\bar{b})$ is the set of all $\bar{a} \in A(\bar{b})$ such that $\bar{b} \in B(\bar{a})$. That is if $\bar{a} \in A^I(\bar{b})$ then there is an $\bar{x}$ such that $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. Similarly we define $B^I(\bar{a})$.

(e) $\bar{R} \subset R$ is the set of all optimal inverse triples.∎

We end this chapter with two examples that help to clarify some of the definitions and foreshadow some of the results of chapter 3.

<u>Example 1-11</u>: We consider the following pair of problems:

P(b) find $\bar{a} = \max\{a = x_1+x_2\}$

subject to

$$2x_1+x_2 \leqq b_1$$
$$x_1+3x_2 \leqq b_2$$
$$x_1,x_2 \geqq 0,$$

I(a) find $\bar{b} = \text{Pmin}\begin{Bmatrix} b_1 = 2x_1+x_2 \\ b_2 = x_1+3x_2 \end{Bmatrix}$

subject to

$$x_1+x_2 \geqq a$$
$$x_1,x_2 \geqq 0.$$

This is a particular case of example 1-8. As in that example we have $B = \bar{B} = \{(b_1,b_2) \geq 0\}$, $A = R$, $\bar{A} = R^+$. This is a linear inverse pair of programming problems.

We now wish to find T, R and $\bar{R}$. Figure 1-2 shows a cross-section of T for a held constant at a positive value. The a value is a feasible value of $P(b)$ for all b in the shaded area and is optimal for b's on the boundary. Hence the boundary - the parts of $b_1 = a$, $b_2 = a$, $2b_1+b_2 = a$ that are shown - is a subset of R. The efficient values of $I(a)$ for the given a, are the values of b satisfying $2b_1+b_2 = 5a$, $b_1 \geq a$, $b_2 \geq a$. Thus for $(\bar{a},\bar{b})$ on this line segment $\bar{a} \in A(\bar{b})$ and $\bar{b} \in B(\bar{a})$. Also notice that the set of equations:

$$2x_1+x_2 = b_1$$

$$x_1+3x_2 = b_2$$

$$x_1+x_2 = a$$

is not independent for values on this line segment. In fact:

$$2(2x_1+x_2 = b_1) + (x_1+3x_2 = b_2)$$

$$\text{yields} \quad 5(x_1+x_2 = 2b_1+b_2)$$

$$\text{which is} \quad 5(x_1+x_2 = a).$$

So for (a,b) on $2b_1+b_2 = 5a$, the three equations have a unique solution. The set of points for which the solution is nonnegative is also the part of the line that forms part of the boundary of the cross-section of T. Hence if $(\bar{a},\bar{b})$ is on the line segment, there is an $\bar{x} \geq 0$ such that $(\bar{a},\hat{b},\bar{x})$ is an optimal inverse triple and these $(\bar{a},\bar{b})$ are contained in $\bar{R}$.

For a < 0, the cross-section is the whole nonnegative

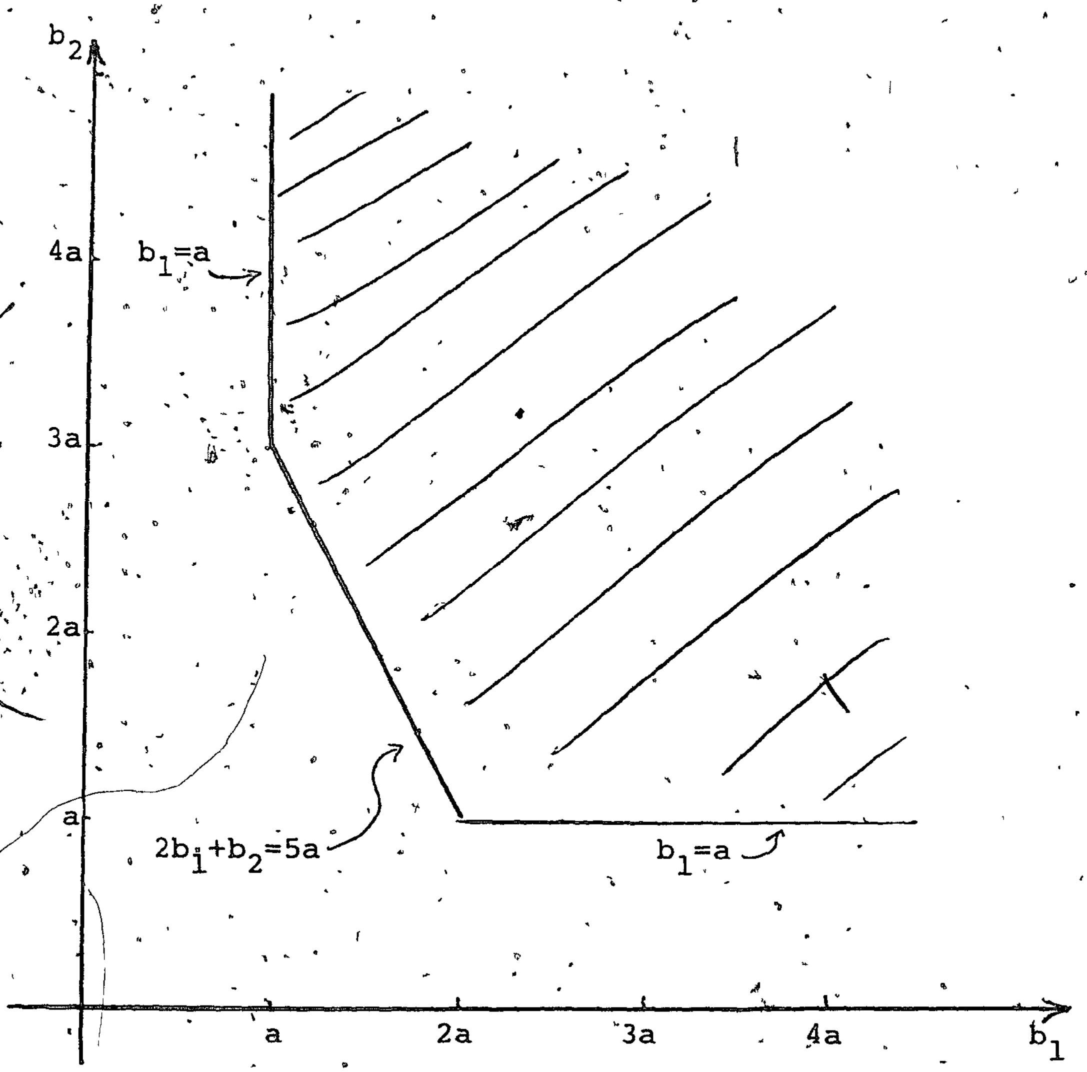
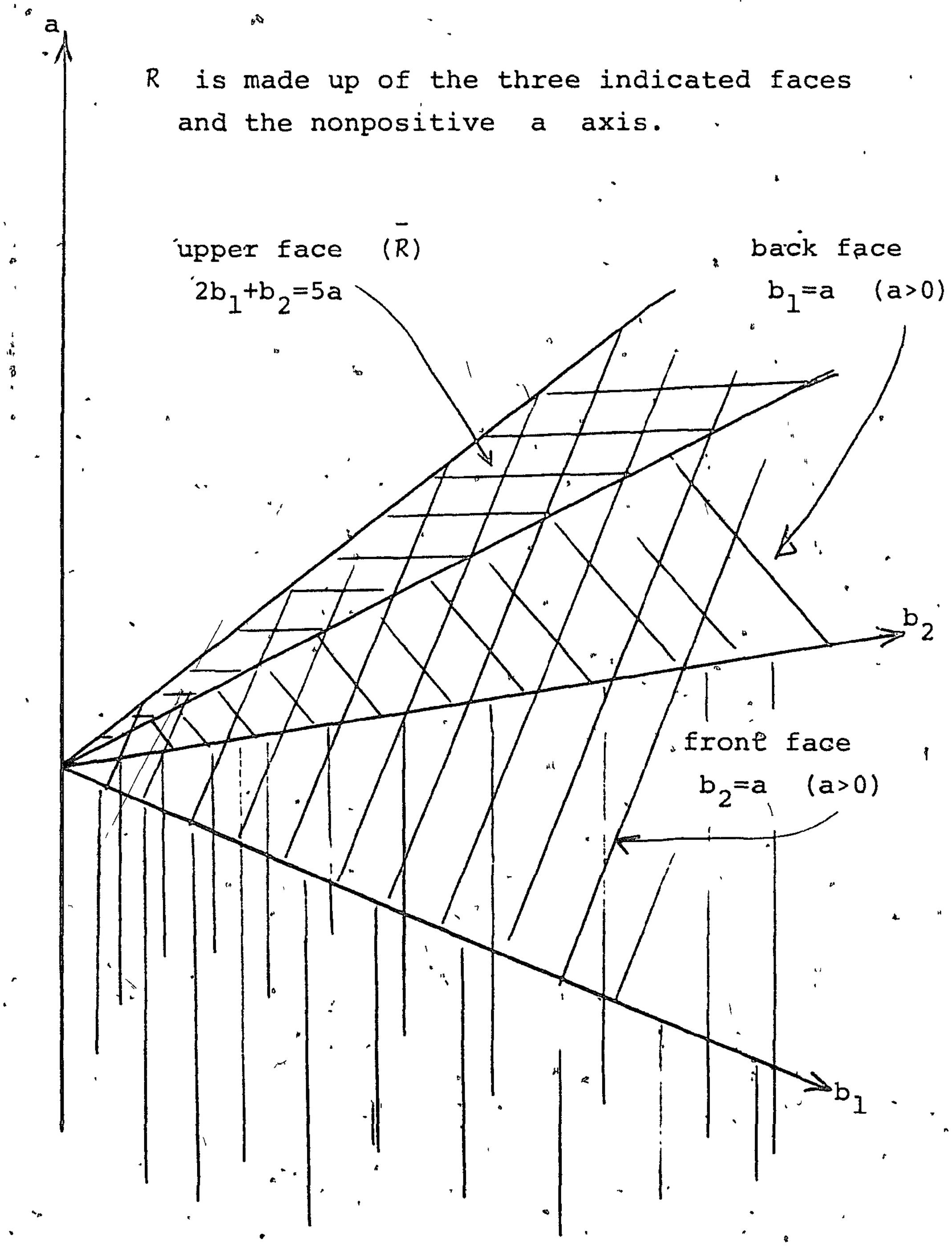


Figure 1-2: Cross-section of T in example 1-11 for a constant $a > 0$.



<u>Figure 1-3:</u> $\mathcal{T}$ for example 1-11.

orthant. Since the only efficient value for $I(a)$ is $(0,0)$ if $a < 0$, $(a,0,0)$ is the only element of $\mathcal{R}$. If a is negative we can raise its value without increasing the efficient value of $I(a)$. This does not contradict definition 1-3 since $a \not\in \bar{\mathcal{A}} = R^+$. This again shows why we use $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$ in definition 1-3 instead of $\mathcal{A}$ and $\mathcal{B}$. $a < 0$ is not an efficient value for $P(b)$ for any $b \varepsilon \bar{\mathcal{B}}$ so there is no $(a,b) \varepsilon \bar{\mathcal{R}}$ with $a < 0$.

Figure 1-3 shows the whole $\mathcal{T}$ set for this example. As above, $\bar{\mathcal{R}}$ is the set of (a,b) in $\mathcal{T}$ such that $2b_1+b_2 = 5a$ or $2/5b_1+1/5b_2 = a$. Notice that these multipliers are also the ones needed in writing the objective function of $P(b)$ as a positive linear combination of the constraint functions. This fact is important and is used in the theorems and algorithms of chapter 3.∎

Example 1-12: We consider the following pair of problems:

P(b) find $\bar{a} = \max\{a = x_1-3x_2\}$

subject to $x_1-2x_2 \leqq b_1$

$x_1-x_2 \leqq b_2$

$x_1,x_2 \geqq 0,$

I(a) find $\bar{b} = \operatorname{Pmax}\left\{\begin{matrix} b_1 = x_1-2x_2 \\ b_2 = x_1-x_2 \end{matrix}\right\}$

subject to $x_1-3x_2 \geqq a$

$x_1,x_2 \geqq 0.$

Again we have a linear inverse pair of programming problems if definition 1-3 is satisfied but in this case some coefficients are negative.

Since we do not obtain a verification of the conditions of definition 1-3 directly from example 1-8 as we did in example 1-11, we first prove that we have an inverse pair. Figure 1-4 shows constraint sets for the pair of problems. In $P(b)$, if b_1 and b_2 are both positive, the optimal solution is at $x = (\min\{b_1,b_2\},0)$. If b_1 or b_2 is negative the optimal solution is at $x = (0, -\min\{b_1/2,b_2\})$. In either case, if b_1 and b_2 are both increased, the optimal value increases. In $I(a)$, the only efficient solution is at the point of intersection of $x_1-3x_2 = a$ with the $x_1 \geqq 0$ axis or the $x_2 \geqq 0$ axis. If a is decreased, this point moves so that both b_1 and b_2 can be decreased. Thus the monotonicity conditions of definition 1-3 are satisfied for all $a \in R$ and $b \in R^2$. So $\bar{A} = A = R$, $\bar{B} = B = R^2$ and we have an inverse pair of problems.

If $a \geqq 0$ and $(a,b) \in T$ we must have $b_1 \geqq a$, $b_2 \geqq a$ but otherwise the b's are not restricted. But a is optimal in $P(b = (a,a))$ and $b = (a,a)$ is optimal in $I(a)$ so $(a,a,a) \in R$. $(a,a,a) \in \bar{R}$ since $(a,(a,a),(a,0))$ is an optimal inverse triple. If $a < 0$ and $(a,b) \in T$ we must have $b_1 \geqq 2/3a$, $b_2 \geqq 1/3a$ but otherwise the b's are not restricted. This time $(a,2/3a,1/3a)$ is in R and $\bar{R}$.

Figure 1-5 shows the complete T set. In this case $R = \bar{R}$. Notice that this time $\bar{R}$ is made up of line segments and also that we can not write x_1-3x_2 (the objective

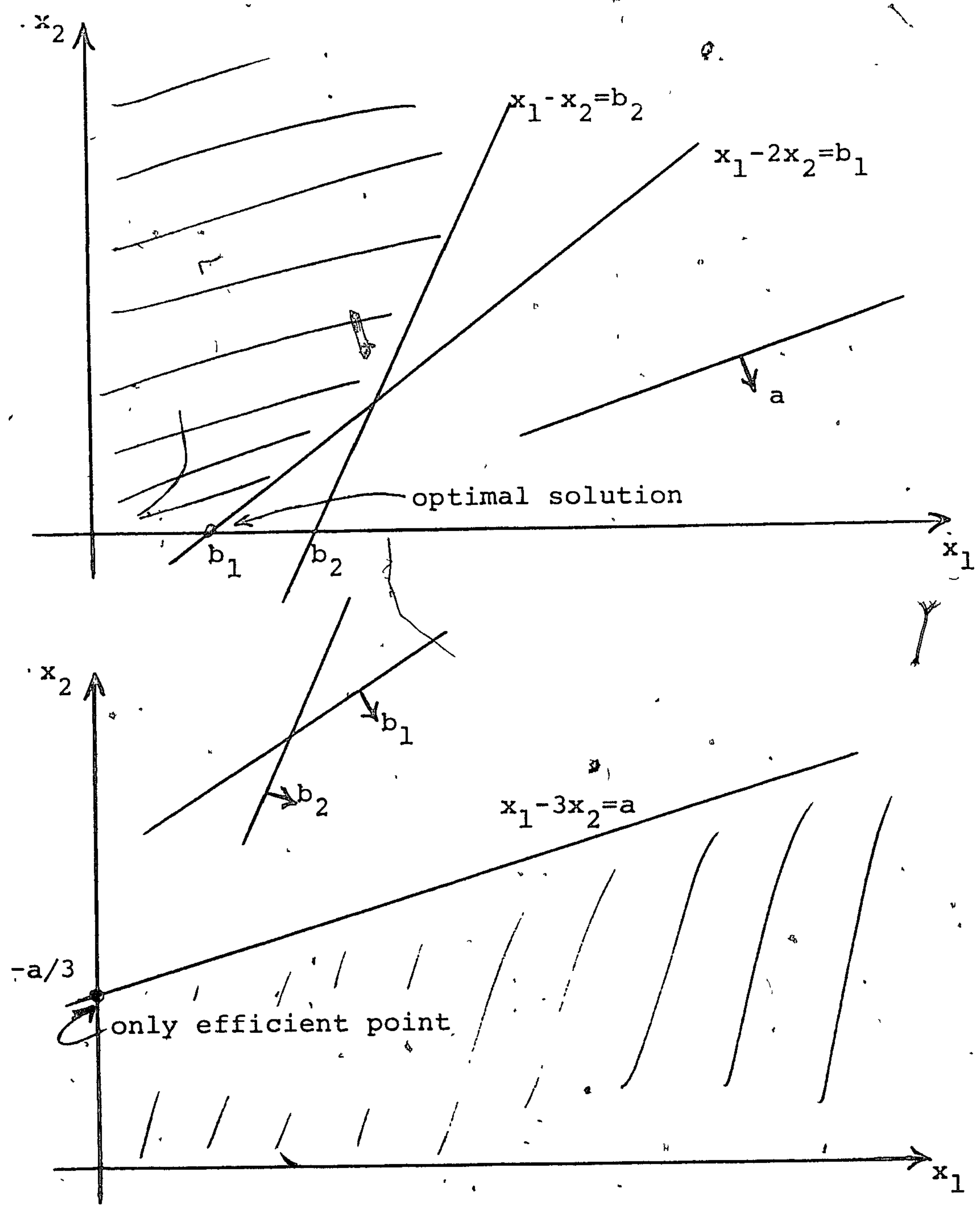


Figure 1-4: Constraint sets for P(b) and I(a) for example 1-12 showing direction of increase of objective functions.

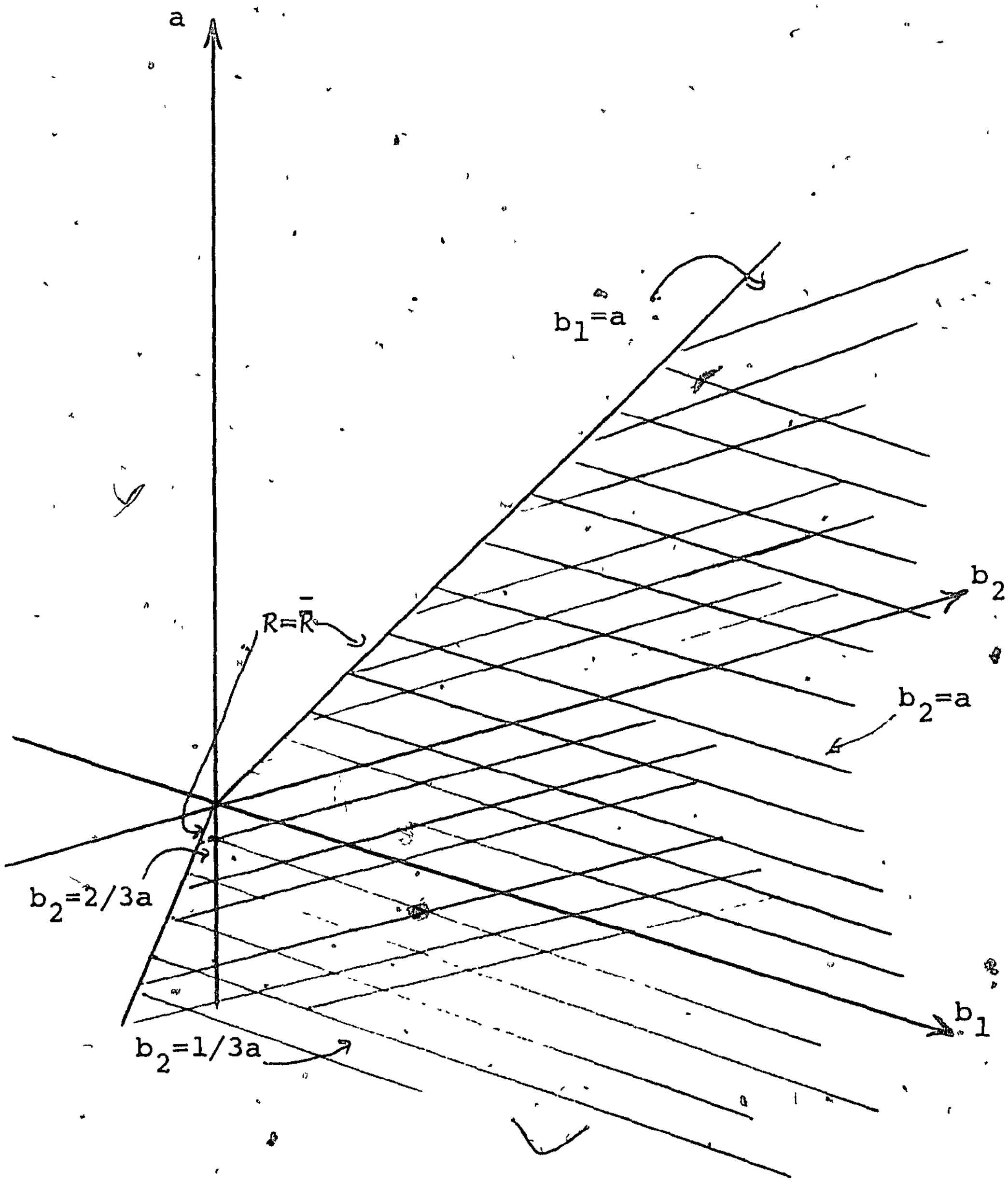


Figure 1-5: T for example 1-12.

function of $P(b)$) as a positive linear combination of x_1-2x_2 and x_1-x_2 (the constraint functions of $P(b)$). This is the converse of what happened in example 1-11 and again is important in chapter 3.∎

Chapter Two - Vector Optimization

In chapter one it is noted that problems with vector objectives arise quite naturally when we consider inverse pairs of problems. In this chapter we review some of the results presently available in vector optimization. In section 2.1 we consider the problem in general and in section 2.2 we discuss what happens to the results when the objectives and constraints are linear.

2.1 General Results

Vector-valued problems arise in situations in which we cannot combine all our objectives into one function. A function that maps the values of several objectives into one number to be optimized is called a utility function. Considerations of the properties of such functions and of when they exist are the subject of utility theory. We do not enter this area of research but just assume that we are dealing with problems for which no utility function exists. If one did we could simply find those points that maximize utility. Otherwise, we usually take the set of Pareto optimal points as the best approximation to the set of utility optimizing points. It is assumed that a utility function cannot be optimized at a point that is not efficient. Therefore most methods for handling vector-valued optimization problems concentrate on ways to find efficient points. The first results we discuss are directed to this end. Later we look at possible methods of reducing the size of the set of efficient points to provide a better approximation to the set of utility optimizing points.

The first discussion of methods to find efficient points was in a paper on nonlinear programming by Kuhn and Tucker in 1950 [21]. Little time seems to have been spent on this problem for the next 15 or so years (see Charnes and Cooper [6] and Karlin [16]) until Geoffrion's paper in 1968 [14]. In the seventies there has been much more work in this area as seen for example in [2,7,11,15,22,25,30,31,32].

One important fact, that appears in most of the early papers, is given in theorem 2-3. First let us define a general vector optimization problem (we use a maximization problem without loss of generality).

<u>Definition 2-1</u>: We consider the following <u>vector maximization</u> problem:

$$V \qquad \text{find } \operatorname{Pmax}\{f(x)=(f_1(x),f_2(x),\ldots,f_p(x))\}$$
$$\text{such that } x \in X, \text{the feasible set.} \blacksquare$$

<u>Definition 2-2</u>: For any $\lambda = (\lambda_1,\ldots,\lambda_p)$ we can form a new <u>scalar-valued problem</u> from V as follows:

$$S_\lambda \qquad \text{find } \max \sum_{i=1}^{p} \lambda_i f_i(x)$$
$$\text{such that } x \in X. \blacksquare$$

<u>Theorem 2-3</u>: Let $\lambda_i > 0$ $(i=1,\ldots,p)$ be fixed. If $\bar{x}$ is

optimal in S_λ, then $\bar{x}$ is efficient in V.

Proof:

∎ This is theorem 1 in Geoffrion [14].∎

However a major problem remains. This does not give all efficient points as the following example shows.

Example 2-4: Let $f(x) = x = (x_1, x_2)$ and and let $X = \{(x_1,x_2) \mid x_1^2 + x_2^2 \leqq 1\}$. Therefore the problem V is to find the Pareto maxima of $\{(x_1,x_2) \mid x_1^2 + x_2^2 \leqq 1\}$. Obviously the set of efficient points is the set

$$E = \{(x_1,x_2) \mid x_1 \geqq 0,\ x_2 \geqq 0,\ x_1^2 + x_2^2 = 1\}$$

as shown in figure 2-1. However the points (1,0) and (0,1) cannot be found by solving a scalar-valued problem with $\lambda_1 > 0$, $\lambda_2 > 0$.∎

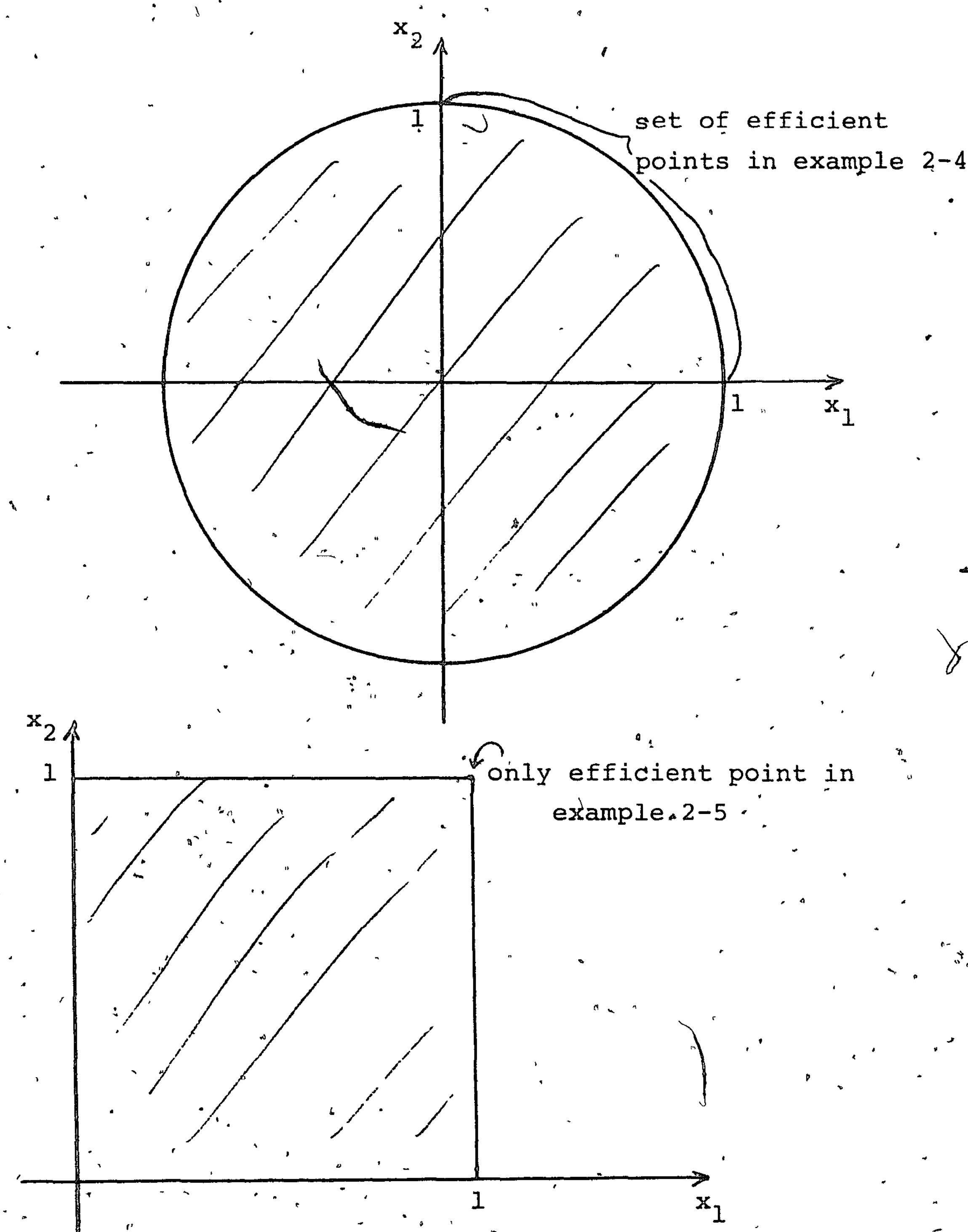
The answer seems to be to only restrict λ_i to be nonnegative. However this results in the following problem.

Example 2-5: Let $f(x) = x = (x_1,x_2)$, and let $X = \{(x_1,x_2) \mid 0 \leqq x_1 \leqq 1,\ 0 \leqq x_2 \leqq 1\}$. V is now:

$$\text{find } \operatorname{Pmax}\{(x_1,x_2) \mid 0 \leqq x_1 \leqq 1,\ 0 \leqq x_2 \leqq 1\}.$$

Now the set of efficient points, E, is just $\{(1,1)\}$. (See figure 2-1.) However if we let $\lambda_1 = 1$ and $\lambda_2 = 0$ we are indifferent to the points $\{(1,x_2) \mid 0 \leqq x_2 \leqq 1\}$. All these are solutions to $S_{(1,0)}$ and only one point is actually in E. ∎

So we are left with the following dilemma. With λ



<u>Figure 2-1</u>: Feasible sets for examples 2-4 and 2-5 showing the sets of efficient points.

positive we do not get all efficient points (even when x is convex and the objectives are linear). Whereas if we allow λ to be just nonnegative we may get too many points. We get only the set of efficient points if the solution to S_λ is unique whenever $\lambda = e_i$ where e_i is the standard unit vector.

In [21] and [14] the authors show that these points occur where there is a first degree gain in one component of the objective for a second degree loss in the others. To get around this problem Geoffrion introduces the notion of proper efficiency defined as follows.

Definition 2-6: x is said to be a properly efficient solution of V if it is efficient and if there exists a scalar $M > 0$ such that, for each i, we have

$$\frac{f_i(x) - f_i(\bar{x})}{f_j(\bar{x}) - f_j(x)} \leq M \text{ for some } j \text{ such}$$

that $f_j(x) < f_j(\bar{x})$ whenever $x \in X$ and $f_i(x) > f_i(\bar{x})$. Geoffrion [14] p.619.■

Thus $\bar{x}$ is not properly efficient if for an arbitrarily large M, there is an $x \in X$ such that the improvement in some component i is at least M times the loss in every other component. But M is arbitrary and there are only a finite set of components. Hence we can make the gain in component i arbitrarily large relative to the loss in any other component.

Example 2-7: Consider example 2-4 again and the point $\bar{x} = (1,0)$. $f(\bar{x}) = \bar{x} = (1,0)$. If x is efficient then $f(x) = x = (\sqrt{1-x_2^2}, x_2)$. Consider the ratio

$$\frac{f_2(x) - f_2(\bar{x})}{f_1(\bar{x}) - f_1(x)} = \frac{x_2 - 0}{1 - \sqrt{1-x_2^2}}$$

Since $\lim\limits_{x_2 \downarrow 0} \dfrac{x_2}{1 - \sqrt{1-x_2^2}} = \lim\limits_{x_2 \downarrow 0} \dfrac{\sqrt{1-x_2^2}}{x_2} = \infty$,

$\dfrac{f_2(x) - f_2(\bar{x})}{f_1(\bar{x}) - f_1(x)}$ cannot be bounded for $x \in X$.

In this example $(1,0)$ and $(0,1)$ are not proper efficient points.■

We see then that if we look for proper efficient points instead of efficient points we are only leaving out undesirable cases. The important benefit we get from looking at this restricted set is the following theorem taken from Geoffrion [14].

Theorem 2-8: Let X be a convex set, and let f_i be concave on X. Then x is properly efficient in V if and only if x is optimal in S_λ for some λ with strictly positive components.

Proof:

■ See Geoffrion [14] page 620, theorem 2.■

Thus as long as our problem satisfies the convexity requirements, we have completely characterized the set of properly efficient points. Geoffrion also notes the following result that shows that the set of proper efficient points is a satisfactory replacement for the set of efficient points.

<u>Lemma 2-9</u>: Let E be the set of all efficient points and $\hat{E}$ be the set of all proper efficient points. Let $f(E)$ and $f(\hat{E})$ be their image under f. If f is concave and continuous and X is closed and convex, then

$$f(\hat{E}) \subseteq f(E) \subseteq \overline{f(\hat{E})}$$

where $\overline{}$ denotes closure.

<u>Proof</u>:

See the references in Geoffrion [14].■

Kuhn and Tucker [21] also define a proper efficiency that differs somewhat from the definition above. However every Kuhn-Tucker proper efficient point is proper efficient in our sense and the two are equivalent if we have the convexity required by theorem 2-8, if f_i and all constraints are differentiable and if the Kuhn-Tucker constraint qualification holds (again see Geoffrion [14]).

A major emphasis of recent work on vector-valued problems has been to try and provide a method by which we can move from the efficient set we have now found to a utility optimizing point. A good reference is the Cochrane and Zeleny book [7] which is the proceedings of a conference held on this subject in 1973. Various methods of attacking the problem are reviewed by the paper by MacCrimmon [22] in this book. These methods fall into the two main categories.

The first method depends on the problem setter being able to give, directly or indirectly, the trade-offs he accepts at any particular set of output levels. One example is as follows. We set the λ vector introduced above and get a proper efficient point. Next we find out in which direction the problem setter would like to move. We now find a new proper efficient point in that direction and repeat. This continues until we cannot find a better efficient point. This assumes that, though we do not know the utility function for the problem, we are able to get some local knowledge about it when needed.

The second major approach is to try to reduce the set of proper efficient points by removing points that are obviously inferior. One method, introduced by Yu and Zeleny (see [30-32]), uses what they call a "dominance structure". Briefly the idea is as follows. When we have an efficient point $\bar{x}$, it is better than any point x such that (1) $f(x) \in f(\bar{x}) + C^1$, and (2) $f(x) \neq f(\bar{x})$ where C^1 is the cone $\{(c_1,\ldots,c_p) | c_i \leqq 0\}$ (the nonpositive orthant). When we have an optimal solution $\hat{x}$ of S_λ, it is better than any point x such that (1) $f(x) \in f(\hat{x}) + C^2$, and (2) $f(x) \neq f(\hat{x})$ where C^2 is the cone $\{c = (c_1,\ldots,c_p) | \lambda c < 0\}$. C^1 or C^2 is called the domination cone and these two provide the extreme cases in which we can find only efficient points, or we can look for a particular maximum since we have a set of multipliers. Between C^1 and C^2 we can find an endless variety of cones that, when used in place of C^1 or C^2 above, give us part of the set of efficient points. If the information about the problem allows us to increase the domination cone, we can eliminate members of the set of proper efficient points that are dominated by other points when the new cone is used. We can also have different cones for different points or sets of points.

These two methods are examples of the two different basic approaches to the problem. Other methods combine ideas from both or use the two iteratively (see [22]).

2.2 Efficient Points in Linear Vector Optimization

In this section we consider problems in which the objective and constraint functions are linear. Our main result is that all efficient points are actually properly efficient points.

First we define a linear vector maximization problem. (Again we consider only maximization problems without any loss of generality.)

<u>Definition 2-10</u>: We consider the following <u>linear vector maximization</u> problem:

$$\text{LV} \qquad \text{find } \operatorname{Pmax}\{Cx\} \quad \text{such that } Dx \leqq b, \quad x \geqq 0$$

where $C \in R^{k \times n}$, $D \in R^{m \times n}$; $x \in R^n$, $b \in R^m$.■

In both Evans and Steuer [11] and Isermann [15], it is proved that all efficient points are proper if the constraints are $Dx = b$. Some simple changes to the proof in [15] yield the following result.

Theorem 2-11: If x^0 is an efficient solution of LV then x^0 is a properly efficient solution of LV.■

Before proving this theorem we prove 3 lemmas.

Lemma 2-12: x^0 is an efficient solution of LV if and only if the linear program

$$\begin{array}{lll} \text{LP1} & \text{maximize} & ey \\ & \text{such that} & Dx \leq b \\ & & -Cx + y = -Cx^0 \\ & & x \geq 0,\ y \geq 0 \\ & & y \in R^k,\ e > 0 \text{ fixed} \end{array}$$

has an optimal solution $\hat{x}, \hat{y}$ with $\hat{y} = 0$.

Proof:

■ If x^0 is efficient in LV then there is no $\bar{x}$ such that $\bar{x} \geq 0$, $D\bar{x} \leq b$ and $C\bar{x} \geq Cx^0$ or $C\bar{x}-y = Cx^0$, $y \geq 0$. Hence if $y \geq 0$, the optimal value of LP1 is zero and this is obtained at $\hat{y} = 0$, $\hat{x} = x^0$.

On the other hand, if $(\hat{x},\hat{y})$ is an optimal solution of LP1 and if $\hat{y} = 0$, then the optimal value of LP1 is zero so there is no $\bar{y} \geq 0$ and $\bar{x} \geq 0$ such that $C\bar{x}-y = Cx^0$, $D\bar{x} \leq b$. If there was, $(\bar{x},\bar{y})$ would be feasible with $e\bar{y} > 0$.

Hence there is no $\bar{x} \geq 0$ such that $C\bar{x} \geq Cx^0$ and $D\bar{x} \leq b$. Therefore x^0 is efficient.∎

<u>Lemma 2-13</u>: x^0 is an efficient solution of LV if and only if the linear program

$$\text{LP2} \qquad \text{minimize } (ub - wCx^0)$$
$$\text{such that } uD - wC \geq 0$$
$$w \geq e > 0$$
$$u \geq 0$$
$$w \in R^k,\ u \in R^m.$$

has an optimal solution $(\hat{u},\hat{w})$ with $\hat{u}b-\hat{w}Cx^0 = 0$.

<u>Proof</u>:

∎ LP2 is the dual of LP1. Therefore $(\hat{x},\hat{y})$ is optimal in LP1 if and only if LP2 has an optimal solution $(\hat{u},\hat{w})$ with $e\hat{y} = \hat{u}b-\hat{w}Cx^0$. Thus x^0 is an efficient solution of LV if and only if $\hat{u}b-\hat{w}Cx^0 = e\hat{y} = 0$.∎

<u>Lemma 2-14</u>: x^0 is an efficient solution of LV if and only if there exists $v^0 > 0$, $v^0 \in R^k$ such that x^0 solves the problem

$$\text{LP3} \qquad \text{maximize } v^0Cx$$
$$\text{such that } Dx \leq b$$
$$x \geq 0.$$

<u>Proof</u>:

∎ Assume x^0 solves LP3 amd x^0 is not efficient in LV. Then there is an $\bar{x}$ feasible in LV (and hence feasible in LP3) such that $C\bar{x} \geq Cx^0$. But this implies

$v^0C\bar{x} > v^0Cx^0$, a contradiction. Hence x^0 is efficient in LV.

On the other hand, assume x^0 is efficient in LV. Therefore LP2 has an optimal solution $(\hat{u},\hat{w})$ which satisfies $\hat{u}b = \hat{w}Cx^0$. Hence $\hat{u}$ is an optimal solution of

LP4

$$\begin{aligned} &\text{minimize } ub \\ &\text{such that: } uD \geq \hat{w}C \\ &\qquad u \geq 0. \end{aligned}$$

Hence an optimal solution exists for the linear program dual to LP4

LP5

$$\begin{aligned} &\text{maximize } \hat{w}Cx \\ &\text{such that } Dx \leq b \\ &\qquad x \geq 0. \end{aligned}$$

From $\hat{u}b = \hat{w}Cx^0$ it follows that x^0 is optimal in LP5. With $v^0 = \hat{w} \geq e > 0$ we obtain x^0 as an optimal solution to LP3.■

<u>Proof of Theorem 2-11</u>:

■ Let x^0 be an efficient solution of LV. Therefore there exists a positive vector $v^0 \in R^k$ such that $v^0Cx^0 \geq v^0Cx$ for all x feasible in LV. By theorem 2-8, x^0 is therefore properly efficient in LV.■

This theorem shows that we can find all efficient solutions to LV by studying solutions to the problem:

$$L(\lambda) \qquad \text{maximize} \quad \lambda Cx$$
$$\text{such that} \quad Dx \leqq b$$
$$x \geqq 0$$
$$\lambda > 0,\ \lambda \in R^k \quad \text{given}$$

as λ is varied parametrically. Usually we normalize λ so that $\sum_{i=1}^{k} \lambda_i = 1$. This avoids repetition of λ's differing only by a multiplicative constant. Since only the slope of the new composite objective function is important, studying $L(\lambda)$ with the normalized λ's gives all possible solutions to the original problem LV. Notice that if LV is infeasible so is each $L(\lambda)$ and so we know after solving $L(\lambda)$ the first time whether or not LV is feasible.

Although the above method appears to solve the problem of how to find the set of all efficient points two problems exist. The first is that solving $L(\lambda)$ for many different choices of λ is long and tedious. This problem can be overcome by applying some of the results of linear programming theory. Using post-optimality analysis we can obtain ways of finding the set of all efficient points once one such point is known. Thus we pick one λ, solve $L(\lambda)$ and then use this solution to help find other efficient points. But one problem still remains. Consider the following example.

Example 2-15: Find $\text{Pmax}\begin{Bmatrix} x_2 \\ -x_1+x_2 \end{Bmatrix}$

LV subject to $-x_1+2x_2 \leqq 2$

$$x_1, x_2 \geqq 0.$$

Now all points on the portion of the line $-x_1+2x_2 = 2$ in the first quadrant are efficient. Consider $L(\lambda)$ with $\lambda = (2/3,1/3)$. (See figure 2-2.)

$L(2/3,1/3)$ maximize $\{2/3x_2 + 1/3(-x_1+x_2)$

$$= -1/3x_1 + x_2\}$$

subject to $-x_1 + 2x_2 \leqq 2$

$$x_1, x_2 \geqq 0.$$

Even though there are finite efficient solutions to the original problem, the new composite objective function is unbounded on the feasible set. Proceeding along the segment of the line $-x_1+2x_2 = 2$ in the first quadrant we are trading off the value of one component of the objective for the value of the other. However, with the multipliers $(2/3,1/3)$ this trade off increases the value of the objective in $L(2/3,1/3)$. Therefore $L(2/3,1/3)$ is not optimized at any of the finite solutions to LV. It turns out in fact that $L(\lambda)$ is unbounded for all normalized λ with $\lambda_1 > 1/2$. If $\lambda_1 < 1/2$, the solution is the efficient point $(0,1)$. Only when $\lambda_1 = 1/2$, are all other efficient points in LV optimal in $L(\lambda)$.∎

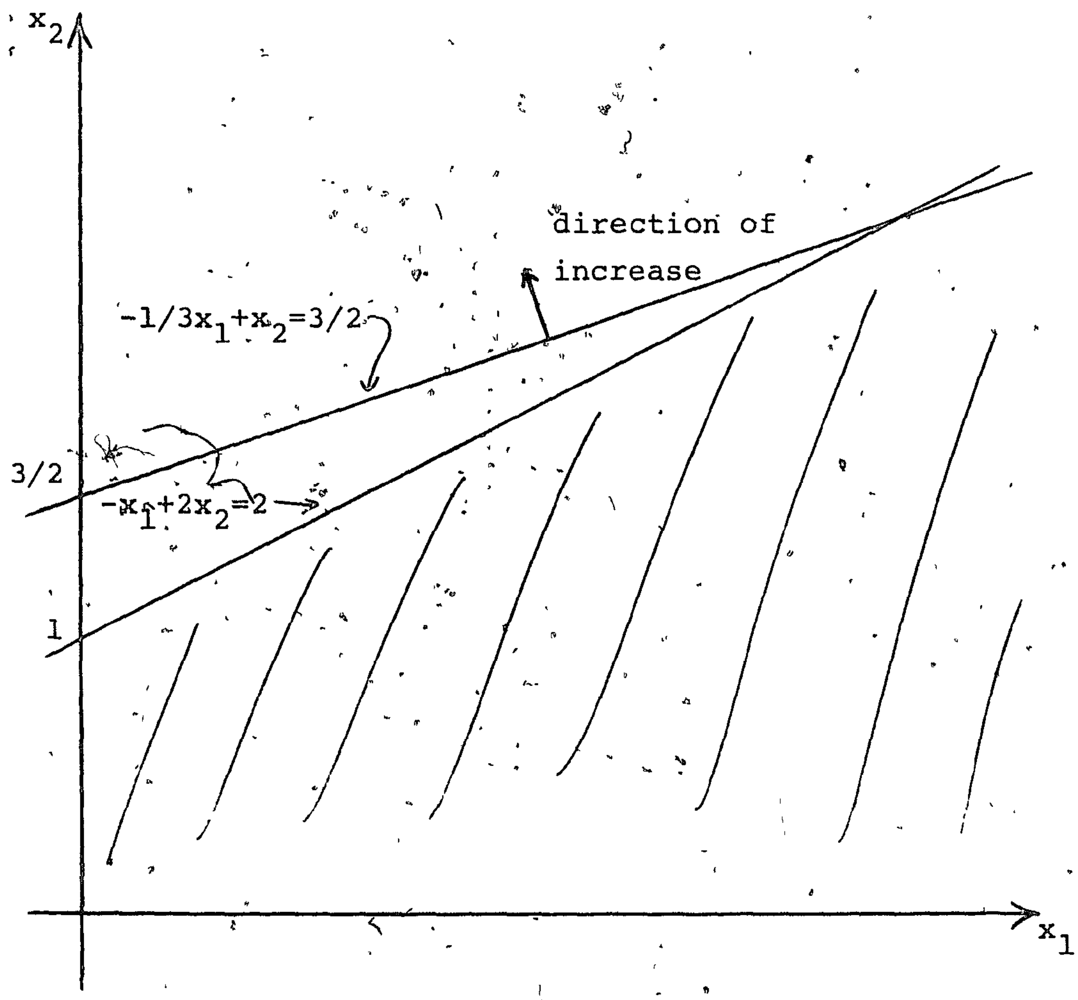


Figure 2-2: Constraint set for example 2-15, showing objective function with $\lambda = (2/3,1/3)$ and with a value of 3/2.

In this example there are finite efficient points, but for the λ picked, the problem $L(\lambda)$ is unbounded.

It can be shown [11] that if LV has at least one finite efficient point, then a finite extreme point must be efficient. In solving $L(\lambda)$ by the simplex method we may pass through an extreme point that is efficient in LV but not optimal in $L(\lambda)$. This is of particular concern if the solution of $L(\lambda)$ is unbounded. Evans and Steuer [11] have a revised simplex method that moves from extreme point to extreme point and can test each or any for efficiency as it progresses. After finding one extreme point their method checks surrounding extreme points for efficiency and in this way builds up the complete efficient set. This procedure works since the set of all efficient points is connected (see [32]). Other algorithms of interest are given in [25] and [32].

<u>Example 2-15, cont'd</u>: The first simplex iteration in the solution of $L(2/3,1/3)$ takes us from the origin to the corner $x = (0,1)$. If we check this point we find it is efficient. Finding the next extreme point to be infinite we get the set E of efficient points of LV to be
$E = \{(x_1,x_2) \mid -x_1+2x_2 = 2;\ x_1 \geq 0,\ x_2 \geq 0\}$. ■

Chapter Three - Linear Inverse Pairs

In this chapter we look at linear inverse pairs of programming problems. First we consider pairs in which one problem is scalar-valued. With insights developed in that case we then consider pairs with vector-valued objectives on both problems. Our main aims are to develop algorithms for finding optimal inverse triples and to show the relationship between an inverse pair and the duals of both problems.

3.1 Scalar-Vector Linear Inverse Problems

We now consider the pair of problems of example 1-8, the case in which one of the problems is a scalar-valued LP problem. Most of the results for this case follow directly from LP theory. Most are also easy to see graphically especially when we have just two variables. However the methods we develop and become familiar with for this case are also useful for problems in which both the primal and the inverse are vector-valued. Our inverse pair can be defined as:

Definition 3-1: The primal problem is

PSL(b) $\quad$ find $\bar{a} = \max\{a = cx\}$

$$\text{such that } Dx \leq b$$
$$x \geq 0$$

The inverse problem is

IVL(a) find $\bar{b} = \text{Pmin}\{b = Dx\}$

such that $cx \geq a$

$x \geq 0$

In both problems $D \in R^{m \times n}$; $c,x \in R^n$; $b \in R^m$; $a \in R$. ∎

PSL(b) is just an LP problem and thus can be solved in the usual way. IVL(a) is a vector-valued problem and can be approached by the methods discussed in chapter 2. However IVL(a) has only one constraint so we can simplify the solution procedure. For ease of discussion we assume $m = 2$, $n = 2$ and the set of efficient points for IVL(a) is non-empty.

Our main concerns in this section are developing algorithms for finding optimal inverse triples and studying their properties. In so doing we have to find solutions to both the primal and inverse problems. The structure helps us to obtain these solutions easily and we examine how and why this is true.

First let us reconsider the solution method for the vector problem IVL(a). We let IVL(a,λ) be the scalar-valued version with multiplier vector λ, that is IVL(a,λ) is the problem:

$$\text{minimize} \quad \lambda b' = \lambda Dx$$

$$\text{such that } cx \geq a$$

$$x \geq 0.$$

In the earlier discussion no importance was put on the actual value of the multipliers. However consider the following example:

Example 3-2: This is a continuation of example 1-11. There we had:

IVL(a) Find $\bar{b} = \text{Pmin}\begin{Bmatrix} b_1 = 2x_1+x_2 \\ b_2 = x_1+3x_2 \end{Bmatrix}$

$$\text{such that } x_1 + x_2 \geqq a$$
$$x_1, x_2 \geqq 0$$

In this simple example it is obvious that the set E of efficient solutions is $E = \{(x_1,x_2) \mid x_1+x_2 = a,\ x_1 \geqq 0,\ x_2 \geqq 0\}$. However, for any normalized multiplier vector with $\lambda_1 < 2/3$, the unique optimal solution of IVL(a,λ) is $(0,a)$ while with $\lambda_1 > 2/3$ the unique optimal solution is $(a,0)$. See figure 3-1. The multiplier vector $(2/3,1/3)$ yields the problem.

$$\text{minimize } \{2/3(2x_1+x_2) + 1/3(x_1+3x_2) = 5/3x_1 + 5/3x_2\}$$

$$\text{subject to } x_1 + x_2 \geqq a$$
$$x_1, x_2 \geqq 0$$

The objective function has the same slope as the single constraint, so the set of all optimal solutions is equal to the set E. It appears that we should consider $(2/3,1/3)$ a better multiplier vector than the others since from it we can easily derive the whole efficient set for IVL(a).

In example 1-11, we show that $\bar{R}$ is the face of T that is also a subset (subcone) of $5a = 2b_1 + b_2$. This

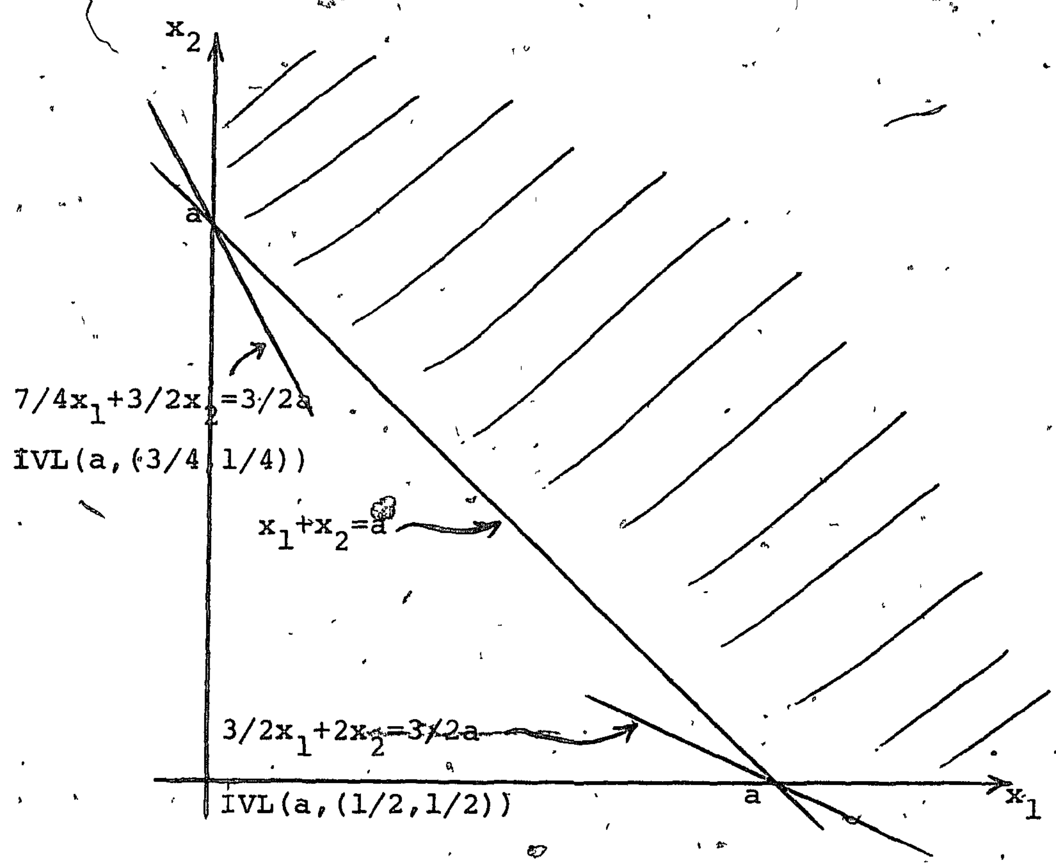


Figure 3-1: Constraint set for example 3-2 showing optimal objective functions for IVL(a,λ) when λ is $(1/2,1/2)$ and $(3/4,1/4)$.

equation can be rewritten as

$$5/3\ a = 2/3\ b_1 + 1/3\ b_2$$

so that the multiplier vector $\bar{\lambda}$ that allows us to find the whole efficient set is made up of the components of the normal to $\bar{R}$ that correspond to b_1 and b_2. As noticed before we also have

$$5/3(1,1) = 2/3(2,1) + 1/3(1,3).$$

Thus $\bar{\lambda}$ is such that for some $\alpha > 0$

$$\alpha c = \bar{\lambda} D. \blacksquare$$

Thus IVL(a) has the following properties. The set of efficient solutions is, depending on the slopes of the objective functions and the constraint, either $\{(x_1,x_2)\,|\,cx = a,\ x_1,x_2 \geq 0.\}$, $\{(\bar{x}_1,0)$ where $\bar{x}_1 = a/c_1\}$ or $\{(0,\bar{x}_2)$ where $\bar{x}_2 = a/c_2\}$. If it is the second or third, then the solution of IVL(a,λ) for any $\lambda > 0$ provides the efficient point. If it is the first, then the solution of IVL(a,λ) for most λ's produces either $(0,\bar{x}_2)$ or $(\bar{x}_1,0)$ and only one normalized λ gives a problem that yields the rest of the efficient set. It would therefore be useful if we could find this "best" λ, call it $\bar{\lambda}$, in any particular case. One answer is provided in [32] and is demonstrated in the example, that is $\bar{\lambda}D = \alpha c$ for some positive constant α. This is the usual equality of slopes condition from Lagrange multiplier theory. If there is no solution to this set of equations then the efficient set is $\{(\bar{x}_1,0)\}$ or

$\{(0,x_2)\}$ and any $\lambda > 0$ suffices to find it.

In this case, if we solve $IVL(a,\lambda)$ with an arbitrarily chosen λ, we most likely end up at one end of the efficient set. But, as we note in chapter 2, the efficient set can usually be taken as an approximation to the set of optimal points under some unknown utility function connecting the objective functions. If we stop with the one efficient point at the edge of the efficient set, we probably have a bad approximation to an optimal solution. In the context of example 1-8, it would appear to be better to pick a solution that outputs some x_1 and x_2 than to take either of the extreme cases of $(\bar{x}_1,0)$ or $(0,\bar{x}_2)$. This is another reason for wanting to find $\bar{\lambda}$.

There is another way of characterizing $\bar{\lambda}$, but first we need the following theorems.

<u>Theorem 3-3</u>: Let $PSL(b)$ and $IVL(a)$ be an inverse pair of problems. If $\bar{x}$ is an efficient solution for $IVL(\bar{a})$ such that $c\bar{x} = \bar{a}$, and $\bar{b} = D\bar{x}$ is an efficient value, then $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple.

<u>Proof:</u>

■ Assume $\bar{a}$ is not optimal in $PSL(\bar{b})$. Then there exist $\hat{a}$ and $\hat{x}$ such that $c\hat{x} = \hat{a} > \bar{a}$, $D\hat{x} \leq \bar{b}$, $\hat{x} \geq 0$. Therefore $\hat{x}$ is feasible in $IVL(\hat{a})$ so there is a $\hat{b} \in B(\hat{a})$ such that $\hat{b} \leq \bar{b}$. But since $\bar{a} < \hat{a}$ and $\hat{b} \in B(\hat{a})$, there is by definition 1-3 a $\bar{\bar{b}} \in B(\bar{a})$ such that $\bar{\bar{b}} < \hat{b}$. We now have $\bar{\bar{b}} < \hat{b} \leq \bar{b}$ or $\bar{\bar{b}} < \bar{b}$. Therefore $\bar{b} \notin B(\bar{a})$ but we have assumed that it is. Hence we have a contradiction. Therefore $\bar{a}$ must be optimal in PSL($\bar{b}$). But then $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple.■

This result depends heavily on cx being scalar and the fact that for scalars $\leq$ and $<$ are the same. We cannot in general obtain the same result if we start with a solution to PSL(b) but the next theorem gives a necessary and sufficient condition for those cases in which the result is valid.

<u>Theorem 3-4</u>: Let PSL(b) and IVL(a) be an inverse pair of problems. If $\bar{x}$ is an efficient solution for PSL($\bar{b}$) and $\bar{a} = c\bar{x}$ is an efficient value, then $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple if and only if the dual of PSL($\bar{b}$) has a positive optimal solution $\bar{y}$.

<u>Proof</u>:

■ First assume $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. Therefore $\bar{a} = c\bar{x}$ and $\bar{b} = D\bar{x}$ and there is no $x \geq 0$ such that either $cx > a$; $Dx \leq b$ or $Dx \leq b$, $cx \geq a$. Since $y\bar{b} \geq c\bar{x}$ holds for any feasible solution to the dual, we require a solution $y > 0$ such that $y\bar{b} \leq c\bar{x}$. Therefore we we need a solution to the system:

(1) $$\begin{aligned} yD &\geq c \\ y &> 0 \\ yb &\leq cx. \end{aligned}$$

We can rewrite this system as:

$$(1') \qquad (\tilde{y},\alpha)I > 0$$

$$(\tilde{y},\alpha)\begin{pmatrix} D \\ -c \end{pmatrix} \geqq 0$$

$$(\tilde{y},\alpha)\begin{pmatrix} -\bar{b} \\ c\bar{x} \end{pmatrix} \geqq 0$$

where $\tilde{y} = \alpha y$, α a scalar. By Motzkin's theorem of the alternative (p.28 in [23]) either (1′) has a solution or (2) has a solution but not both, where (2) is the system:

$$(2) \qquad Iz_1 + \begin{pmatrix} D \\ -c \end{pmatrix} z_2 + \begin{pmatrix} -\bar{b} \\ c\bar{x} \end{pmatrix} \theta = 0$$

$$z_1 \geq 0,\ z_2 \geqq 0,\ \theta \geqq 0$$

$$z_1 \in R^{m+1},\ z_2 \in R^n,\ \theta \in R.$$

This system may be rewritten as:

$$(2') \qquad Dz_2 + z_1' = \bar{b}\theta$$

$$cz_2 - \eta = c\bar{x}\theta$$

$$z_1 = (z_1', \eta) \geq 0,\ z_2 \geqq 0,\ \theta \geqq 0$$

$$z_1' \in R^m,\ \eta \in R,\ z_2 \in R^n,\ \theta \in R.$$

Since $z_1 \geq 0$ either $z_1' \geq 0$ and $\eta \geqq 0$ or $z_1' \geqq 0$ and $\eta > 0$. Therefore, (2′) has a solution if and only if one of the following two systems, (2-1) and (2-2), has a solution:

$$(2\text{-}1) \qquad cz_2 > c\bar{x}\theta$$

$$Dz_2 \leqq \bar{b}\theta$$

$$z_2 \geqq 0,\ \theta \geqq 0;$$

$$(2\text{-}2) \qquad cz_2 \geqq c\bar{x}\theta$$

$$Dz_2 \leq b\theta$$

$$z_2 \geqq 0,\ \theta \geqq 0.$$

First we assume $\theta > 0$ and without loss of generality we can actually assume $\theta = 1$. Now (2-1) becomes:

$$cz_2 > c\bar{x} ,$$

$$Dz_2 \leqq \bar{b}$$

$$z_2 \geqq 0.$$

But this system cannot have a solution since $\bar{x}$ is optimal in $PSL(\bar{b})$. (2-2) becomes:

$$Dz_2 \leq b$$

$$cz_2 \geqq cx$$

$$z_2 \geqq 0$$

Again this system cannot have a solution since $\bar{x}$ is optimal in $IVL(\bar{a})$. Hence if (2) has a solution θ must equal zero. In this case (2-1) becomes:

$$cz_2 > 0$$

$$Dz_2 \leqq 0$$

$$z_2 \geqq 0.$$

Now for any y, feasible in the dual of $PSL(\bar{b})$, $yD \geqq c$ so, combining this with the above, we have:

$$0 \geqq yDz_2 \geqq cz_2.$$

But we have $cz_2 > 0$. Hence, if the dual has a feasible solution there can be no solution to (2-1). For (2-2) we get:

$$cz_2 \geqq 0$$

$$Dz_2 \leq 0$$

$$z_2 \geqq 0.$$

Consider now $\bar{x} + z_2$. We have $D(\bar{x} + z_2) = D\bar{x} + Dz_2 \leq b + 0 = b$; $c(\bar{x} + z_2) = c\bar{x} + cz_2 \geqq c\bar{x} + 0 = c\bar{x}$ and $\bar{x} + z_2 \geqq 0$. Therefore, if z_2 is a solution to this system, $\bar{x} + z_2$ solves (2-2) for $\theta = 1$. But there was no solution when $\theta = 1$, so there can be no solution z_2 to this system. Therefore we conclude that (2) has no solution so (1) must have one. Hence if $(\bar{a}, \bar{b}, \bar{x})$ is an optimal inverse triple we can find a positive optimal solution to the dual of $PSL(\bar{b})$.

Now assume that the dual of $PSL(\bar{b})$ has a positive optimal solution $\bar{y}$. So system (1) above has a solution and system (2) cannot. Hence, in particular, there is no $x \geqq 0$ such that either $cx > \bar{a}$, $Dx \leqq \bar{b}$ or $Dx \leqq \bar{b}$, $cx \geqq \bar{a}$. We assumed that $\bar{x}$ is an efficient solution for $PSL(\bar{b})$ and these inequalities imply that $\bar{x}$ is also an efficient solution for $IVL(\bar{a})$. We know $\bar{a} = c\bar{x}$ and, since $\bar{y} > 0$, complementary slackness in the primal problem gives us $\bar{b} = D\bar{x}$. Therefore $(\bar{a}, \bar{b}, \bar{x})$ is an optimal inverse triple.∎

We now are in a position to develop a new characterization of $\bar{\lambda}$. The following theorem gives the theory we need. After that we explain the charaterization as a corollary.

<u>Theorem 3-5</u>: Let $PSL(b)$ and $IVL(a)$ be an inverse pair of problems and let $(\bar{a}, \bar{b}, \bar{x})$ be an optimal inverse triple, then:

(1) if $\bar{y} > 0$ solves the dual of $PSL(\bar{b})$, $\bar{b}$ is an efficient value of $IVL(\bar{a})$ generated by solving

$IVL(\bar{a},\bar{y})$;

(2) if $\bar{b}$ is an efficient value of $IVL(\bar{a})$ generated by solving $IVL(\bar{a},\lambda)$ then $y = \alpha\lambda$ solves the dual of $PSL(\bar{b})$ for some $\alpha > 0$ and, if $\bar{\lambda}\bar{b} \neq 0$, $\alpha = \dfrac{\bar{a}}{\bar{\lambda}\bar{b}}$.

Proof:

Part (1) follows directly from theorem 4-16 and is proved as a corollary to that theorem.

Part (2). Since, by theorem 3-4, the dual of $PSL(\bar{b})$ has a positive optimal solution, it must be feasible. We can write the dual as: $\min\ y\bar{b}$

$$yD \geqq c$$

$$y \geqq 0$$

By assumption, we also know $\bar{a} = c\bar{x}$, $\bar{b} = D\bar{x}$.

First let us assume that $\lambda\bar{b} \neq 0$. If $\dfrac{\bar{a}}{\lambda\bar{b}}\cdot\lambda$ is feasible it must be optimal, since the value of the dual would be $\dfrac{\bar{a}}{\lambda\bar{b}}\cdot\lambda.\bar{b} = \bar{a}$. Therefore, for our result to hold we need to have a solution α to the system:

(2-1)
$$\alpha\lambda D \geqq c$$

$$\alpha = \frac{\bar{a}}{\lambda\bar{b}}$$

$$\alpha \geqq 0$$

But, using a theorem of the alternative due to Gale (p.35 in [23]), we can deduce that there is either a solution to (2-1) or a solution to (2-2) but not both, where (2-2) is the system:

$$\text{(2-2)} \qquad \begin{aligned} \lambda Dz &\leq \beta \\ cz &> \frac{\bar{a}}{\lambda\bar{b}}\,\beta \\ z &\geq 0 \\ z \in R^n&,\ \beta \in R. \end{aligned}$$

Now assume $\beta = 0$ so:

$$\begin{aligned} \lambda Dz &\leq 0 \\ cz &> 0. \end{aligned}$$

If we let $\bar{z} = z + \bar{x}$ we have:

$$c\bar{z} = c\bar{x} + cz > \bar{a} + 0 \quad \text{and}$$

$$\lambda D\bar{z} = \lambda D\bar{x} + \lambda Dz \leq \lambda D\bar{x} = \lambda\bar{b}$$

so we have a solution to:

$$\text{(2-3)} \qquad \begin{aligned} c\bar{z} &> \bar{a} \\ \lambda Dz &\leq \lambda b \\ z &\geq 0 \end{aligned}$$

For $\beta > 0$ we must consider two cases, $\lambda\bar{b} > 0$ or $\lambda\bar{b} < 0$. If $\lambda\bar{b} > 0$, we can rewrite (2-2) as:

$$\begin{aligned} \lambda D\,\frac{\lambda\bar{b}}{\beta}\,z &\leq \lambda\bar{b} \\ c\,\frac{\lambda\bar{b}}{\beta}\,z &> \bar{a} \\ z &\geq 0 \end{aligned}$$

Therefore $\bar{\bar{z}} = \frac{\lambda\bar{b}}{\beta}\,z$ solves (2-3). If $\lambda\bar{b} < 0$ we have:

$$\begin{aligned} c\,\frac{\lambda\bar{b}}{\beta}\,z &< \bar{a} \\ \lambda D\,\frac{\lambda\bar{b}}{\beta}\,z &\geq \lambda\bar{b} \end{aligned}$$

so:

$$c(\bar{x} - \frac{\lambda\bar{b}}{\beta} z) > \bar{a} - \bar{a} = 0$$

$$\lambda D(\bar{x} - \frac{\lambda\bar{b}}{\beta} z) \leqq \lambda\bar{b} - \lambda\bar{b} = 0$$

$$\text{and } \bar{x} - \frac{\lambda\bar{b}}{\beta} z \geqq 0 \quad \text{since } \frac{\lambda\bar{b}}{\beta} z \leqq 0.$$

Therefore, letting $\bar{z} = 2\bar{x} - \frac{\lambda\bar{b}}{\beta} z$, we again get a solution to (2-3). Similarly, if $\beta < 0$, we can also obtain a solution to (2-3). Now let $\hat{a} = c\bar{z}$ so $\hat{a} > \bar{a}$. Consider $IVL(\hat{a})$. $\bar{z}$ is feasible in this problem so it has an optimal value $\hat{b} \leqq D\bar{z}$. Since $\lambda > 0$ we have $\lambda\hat{b} \leqq \lambda D\bar{z}$. Now we can use definition 1-3 and $\{\bar{a} < \hat{a}$ and $\hat{b} \in B(\hat{a})\} \Rightarrow \{\exists \tilde{b} \in B(\bar{a}) | \tilde{b} < \hat{b}\}$. Therefore, putting all these together,

$$\lambda\tilde{b} < \lambda\hat{b} \leqq \lambda D\bar{z} \leqq \lambda\bar{b}$$

$$\text{or} \quad \lambda\tilde{b} < \lambda\bar{b}.$$

However $\tilde{b} \in B(\bar{a})$ so there must be an $\tilde{x} \geqq 0$ such that $\tilde{b} = D\tilde{x}$ and $c\tilde{x} \geqq \bar{a}$. Hence $\tilde{x}$ is feasible in $IVL(\bar{a}, \lambda)$. But $\lambda\bar{b}$ is the optimal value of the problem so there can be no $\tilde{b}$ and therefore there is no solution to (2-3) and hence no solution to (2-2). Therefore there is a solution to (2-1) so $\frac{\bar{a}}{\lambda\bar{b}}\lambda$ is feasible in the dual of $PSL(\bar{b})$ and is therefore an optimal solution. (Notice that if $\bar{a} \neq 0$, then $\bar{a}$ and $\lambda\bar{b}$ must have the same sign since $\lambda > 0$).

Now we turn to the case in which $\lambda\bar{b} = 0$. Therefore $\alpha\lambda\bar{b} = 0$ for all α and we need to show two things. For the dual of $PSL(\bar{b})$, $\alpha\lambda$ is feasible for some $\alpha > 0$, and $y\bar{b} \geqq 0$ for all feasible y. Assume the second part is false. Thus there is a solution to:

$$\text{(2-4)} \qquad \bar{y}\bar{b} < 0$$
$$\bar{y} \geqq 0$$
$$\bar{y}D \geqq c.$$

and since $\bar{x} \geqq 0$,

$$0 > \bar{y}\bar{b} = \bar{y}D\bar{x} \geqq c\bar{x} = \bar{a}$$

so $\bar{a}$ must be negative. But then $x = 0$ is feasible for IVL(0) so there exists some $\hat{b} \in B(0)$ and $\hat{b} \leq 0$. Using definition 1-3, $\{\bar{a} < 0$ and $\tilde{b} \in B(0)\} \Rightarrow \{\exists\tilde{b} \in B(\bar{a}) | \tilde{b} < \hat{b}\}$. Thus $\tilde{b} < 0$, so that $\lambda\tilde{b} < 0$ is a feasible value for IVL$(\bar{a},\lambda)$ which contradicts the optimality of $\lambda\bar{b} = 0$. Hence $\bar{y}\bar{b} \geqq 0$ holds for all y which are feasible in the dual of PSL$(\bar{b})$.

Now if $\alpha\lambda$ is feasible for some $\alpha > 0$, there is a solution to the system

$$\text{(2-5)} \qquad -\alpha\lambda D \leqq -c$$
$$\alpha > 0.$$

But by a theorem of the alternative due to Gale (p.35 in [23]) we can show that (2-5) has a solution only if (2-6) does not, where (2-6) is:

$$\text{(2-6)} \qquad \begin{bmatrix} \lambda D \\ -c \end{bmatrix} z \leq 0$$
$$z \geqq 0.$$

(2-6) can be rewritten as: for some $z \geqq 0$

either $\qquad \lambda Dz < 0$ and $cz \geqq 0$

or $\qquad \lambda Dz \leqq 0$ and $cz > 0.$

Letting $\bar{z} = \bar{x} + z$, $\bar{z}$ must solve either

$$c\bar{z} \geqq \bar{a} \quad \text{and} \quad \lambda D\bar{z} < \lambda\bar{b}$$

or $c\bar{z} > \bar{a}$ and $\lambda D\bar{z} \leq \lambda b$.

In the latter we have a solution to (2-3) which leads to the same contradiction as derived earlier. In the former $\bar{z}$ is feasible in $IVL(\bar{a},\lambda)$ but yields a better value than the optimum. Hence again we get a contradiction. Therefore $\alpha\lambda$ is feasible for some $\alpha > 0$. Therefore the proof is complete.■

<u>Corollary</u>: Let $PSL(b)$ and $IVL(a)$ be an inverse pair of problems and let $BAR = \{\bar{b} \mid \bar{b} = D\bar{x}$ is an efficient value in $IVL(\bar{a})$ and $c\bar{x} = \bar{a}\}$, then there exists a common positive solution, $\bar{y}$, to the dual of each $PSL(\bar{b})$ for all $\bar{b} \in BAR$ and all $\bar{b}$ can be derived as solutions to $IVL(a,\bar{y})$, that is $\bar{y} = \alpha\bar{\lambda}$ for some $\alpha > 0$.

<u>Proof</u>:

■ For each $\bar{b} \in BAR$ and corresponding $\bar{x}$, $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. All the $\bar{b}$ can be generated by solving $IVL(\bar{a},\bar{\lambda})$ and for each $\bar{b}$ there is an $\alpha > 0$ such that $\alpha\bar{\lambda}$ solves the dual of $PSL(\bar{b})$. Hence $\bar{\lambda}\bar{b}$ is constant for all $\bar{b} \in BAR$. If $\bar{\lambda}\bar{b} \neq 0$, $\alpha = \dfrac{\bar{a}}{\bar{\lambda}\bar{b}}$ by theorem 3-5 so α is constant for all $\bar{b}$. If $\bar{\lambda}\bar{b} = 0$, any α for which $\alpha\bar{\lambda}$ solves one of the dual problems suffices.■

Theorem 3-5 is intuitively reasonable since both the dual variable and the multiplier for a $D_i x = b_i$ (where D_i is the i^{th} row of D) measure the value of this component. The dual variable is a derived value when b_i is given. The multiplier is a given value that leads to a derived b_i.

In the above theorem the α is necessary since all positive scalar multiples of a multiplier vector are equivalent, but, if we multiply an optimal set of dual variables by a scalar not equal to 1, the result is either not feasible or not optimal. We have not normalized the multiplier vector in part (1) of the theorem but an equivalent result would be that $\bar{b}$ is efficient in IVL$(\bar{a},\hat{\lambda})$ where $\hat{\lambda} = \bar{y}/\sum_{i=1}^{m} \bar{y}_i$.

We now turn to algorithms for finding optimal inverse triples. In the first we assume an $a = \bar{a}$ is given, and we are looking for all $\bar{b}$ and $\bar{x}$ such that $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple.

Algorithm 3-6: This algorithm finds optimal inverse triples for a specified $\bar{a} \in \bar{A}$ for an inverse pair of problems.

(1) Choose $\lambda > 0$ and solve the simple linear programming problem IVL$(\bar{a},\lambda)$ for $\bar{x}$ and $\bar{b}$.

(2) If $\bar{x} > 0$ then $\lambda = \bar{\lambda}$ and proceed to step 4.

(3) Solve the linear program which is the dual of PSL$(\bar{b})$. If there is no positive solution, $\bar{a} \neq c\bar{x}$ and there are no optimal inverse triples for $\bar{a}$ so stop. If there is no extreme point positive optimal solution, then $\bar{x}$ (from step 1) is the entire efficient set of solutions. Otherwise let $\bar{\lambda}$ be the positive extreme point optimal solution.

(4) Solve the simple linear programming problem IVL$(\bar{a},\bar{\lambda})$ and determine the entire set of optimal inverse triples for this $\bar{a}$.■

In step (2), $\lambda = \bar{\lambda}$ since $\bar{\lambda}$ is the only multiplier vector that does not yield either the $(\bar{x}_1,0)$ or $(0,\bar{x}_2)$ (defined previously) as the only optimal solution in step (1). If we do get $(\bar{x}_1,0)$ or $(0,\bar{x}_2)$ then the dual of PSL($\bar{b}$) must have alternative optimal solutions. If the set of efficient solutions does not contain just one point, $\bar{\lambda}$ must be an extreme point solution to this dual. If no extreme point solution is positive, then the solutions of the dual, when normalized, give all normalized multipliers. Thus the $\bar{x}$ found is the only efficient solution of IVL($\bar{a}$).

Example 3-2, cont'd: Say we first pick $\lambda = (1/2,1/2)$. As we saw before the optimal solution to IVL(10, (1/2,1/2)) is $x = (10,0)$ yielding $\bar{b} = (20,10)$. PSL(20,10) is then:

$$\bar{a} = \max a = x_1+x_2$$
$$\text{such that } 2x_1+x_2 \leqq 20$$
$$x_1+3x_2 \leqq 10$$
$$x_1,x_2 \geqq 0$$

And its dual (see figure 3-2) is:

$$\text{minimize } 20y_1+10y_2$$
$$\text{such that } 2y_1+y_2 \geqq 1$$
$$y_1+3y_2 \geqq 1$$
$$y_1,y_2 \geqq 0$$

This problem has two basic optimal solutions. One of these is positive and if we normalize it we get

$$\bar{\lambda} = \frac{(2/5,1/5)}{(3/5)} = (2/3,1/3).$$ This is the multiplier vector that we showed above yields the whole efficient set.∎

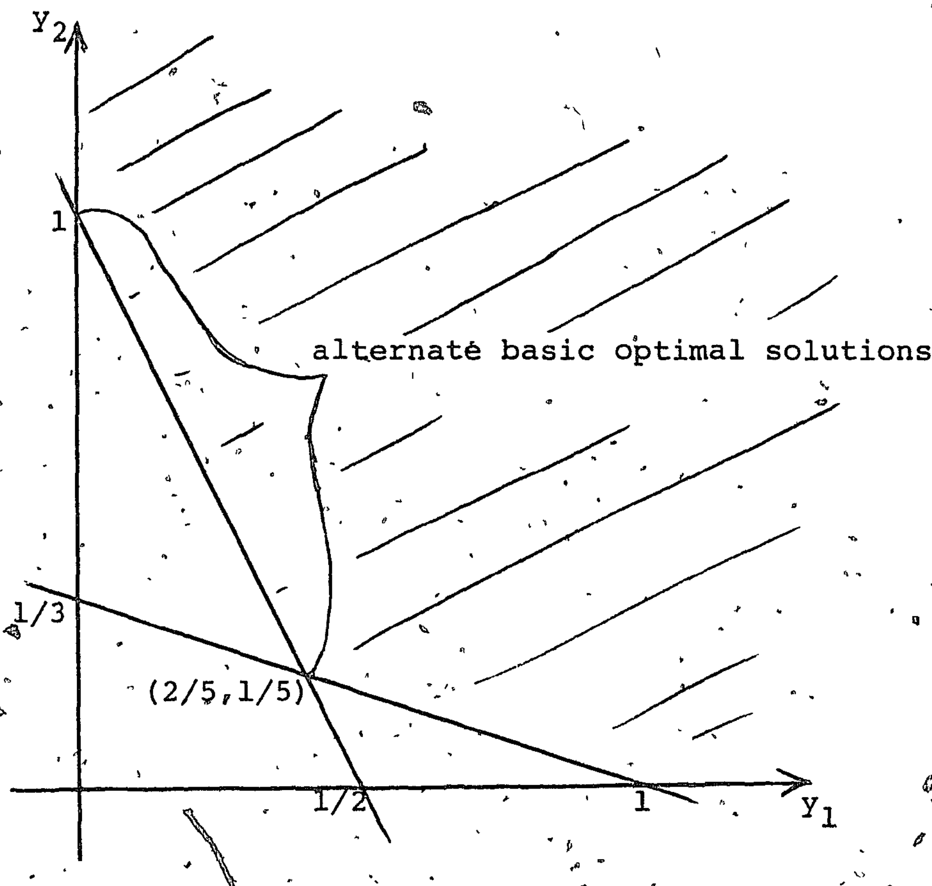


Figure 3-2: Constraint set for the dual of PSL(20,10) in example 3-2 showing alternate basic optimal solutions.

In the next algorithm we again start with a specified value for $\bar{a}$ but this time we want the components of $\bar{b}$ to be in some given ratio. There may be no optimal inverse triple that has $\bar{b}$ with exactly the desired ratio of components and the ratio has to be adjusted. There are many ways of doing this. We do it by resetting the ratio so that, at the optimal solution to the original problem, all constraints are tight.

<u>Algorithm 3-7</u>: This algorithm finds an optimal inverse triple $(\bar{a},\bar{b},\bar{x})$ for an inverse pair of problems in which the value $\bar{a} \in \bar{A}$ is specified and where the vector $\bar{b}$ is proportional to a given $\hat{b}$.

(1) Solve the linear programming problem $PSL(\hat{b})$ for $\hat{x}$ and $\hat{a}$. If $\hat{a} = \bar{a}$, let $\bar{b} = \hat{b}$, $\bar{x} = \hat{x}$ and proceed to step (2). If $\hat{a} = 0 \neq \bar{a}$ proceed to step (5). Otherwise let $\gamma = \bar{a}/\hat{a}$ and note that $\bar{x} = \gamma\hat{x}$ solves $PSL(\bar{b}=\gamma\hat{b})$ with optimal value $\bar{a}$.

(2) If $D\bar{x} \leq \bar{b}$ then reset $\bar{b} = D\bar{x}$. The desired proportionality cannot be exactly achieved.

(3) Solve the linear programming problem that is dual to $PSL(\bar{b})$. If it has a positive optimal solution then $(\bar{a},\bar{b},\bar{x})$ is the desired triple.

(4) Otherwise the problem $PSL(\bar{b})$ has an alternative optimal solution $\bar{\bar{x}}$ for which $c\bar{\bar{x}} = \bar{a}$ and $D\bar{\bar{x}} \lneq \bar{b}$. Set $\bar{x} = \bar{\bar{x}}$ and return to step (2).

(5) Multiply $\hat{b}$ by some positive constant not equal

to 1 and return to (1). If $\hat{a}$ is still 0, the desired proportionality cannot be exactly achieved. Increase all components of $\hat{b}$ slightly and return to (1). If $\hat{a}$ is still 0 and $D\hat{x} \not\leq \hat{b}$ definition 1-3 is violated and optimal inverse triples cannot exist. If $D\hat{x} < \hat{b}$, reset $\hat{b} = D\hat{x}$ and return to (1).∎

Part (3) follows theorem 3-4. What we really need in part (4) is a $\bar{\bar{b}}$ efficient in $IVL(\bar{a})$. That such a $\bar{\bar{b}}$ exists with an $\bar{\bar{x}}$ such that $c\bar{\bar{x}} = \bar{a}$ follows because of the way $\bar{b}$ has been defined, because $\bar{a} \in \bar{A}$ and because PSL(b) and IVL(a) are an inverse pair. Also consider the set of optimal solutions of $PSL(\bar{b})$. There must be an extreme point $\tilde{x}$ of this set such that $c\tilde{x} = \bar{a}$, $D\tilde{x} \leq \bar{b}$ and there is no other x in this set with $Dx \leq D\tilde{x}$ or else $Dx = b$ for all x in the set. This is true since $\tilde{x}$ is just an extreme point efficient solution for the problem $IVL(\bar{a})$. Thus we only need consider extreme points of the original set of optimal solutions. This set is finite, so the algorithm terminates. If all components of $\hat{b}$ are of the same sign, then the first adjustment of $\hat{b}$ in part (5) suffices, since if $\hat{a}$ is zero and $D\hat{x} \not\leq \hat{b}$ we already have a violation of definition 1-3. The extra part of the step allows us to handle cases in which the components of $\hat{b}$ have different signs. This algorithm can also be used to find optimal inverse triples with a $\bar{b}$ close to a given value if the value of $\bar{a}$ is unrestricted. By "close to" we mean $\bar{b}$ is reset as necessary as outlined in the algorithm.

<u>Example 3-2, once again</u>: Take $\bar{a} = 10$ again. Say we guess $\hat{b} = (10,10)$ and solve the problem:

$$\text{PSL}(10,10) \qquad \hat{a} = \{\max a = x_1+x_2\}$$
$$\text{such that } 2x_1+x_2 \leqq 10$$
$$x_1+3x_2 \leqq 10$$
$$x_1,x_2 \geqq 0$$

The optimal solution to PSL(10,10) is $\hat{x} = (4,2)$ with value $\hat{a} = 6$. Therefore $\gamma = \bar{a}/\hat{a} = 10/6 = 5/3$ and PSL(50/3,50/3) has the optimal solution $\bar{x} = (20/3,10/3)$ and value $10 = \bar{a}$. So $T = (10,(50/3,50/3), (20/3,10/3))$ is our candidate for an optimal inverse triple and we check for a positive solution to:

$$\text{minimize } 50/3y_1+50/3y_2$$
$$\text{such that } 2y_1+y_2 \geqq 1$$
$$y_1+3y_2 \geqq 1$$
$$y_1,y_2 \geqq 0$$

This problem has the unique optimal solution $(2/5,1/5)$ so T is the desired triple.∎

In the preceding discussion and examples, we assume $m = 2$ and $n = 2$. However this restriction is not used in the proofs of theorem 3-3, 3-4 or 3-5 or algorithm 3-7, and all these results hold for any m and n. In the above, the efficient set for IVL(a) is a line segment or a point. If there are more than 2 variables the efficient set is the segment of a plane or hyperplane cut off by the positive orthant or a corner or edge of the segment. Algorithm 3-6 must be modified to handle this. If the efficient set has dimension $n - 1$, that is, it is all of the segment cut off by the positive orthant, or is one point the algorithm works. Otherwise we are trying to write c as a positive combination of the rows of D but cannot. Since we have an inverse pair of problems and $\bar{a} \in \bar{A}$ we know optimal inverse triples exist. If definition 1-3 is satisfied, we must be able to write c as a positive combination of the rows of D and some of the rows of $-I$ from the nonnegativity constraints $-Ix \leqq 0$ (see for example [9]). By adding as few rows of $-I$ as possible we find which constraints must be active and thus reduce the efficient set. The following examples demonstrate these points. In the first the efficient set is still easy to see since there are still just two variables but it does indicate what can happen if n and m differ. In the second we consider a case in which m and n are both 3.

<u>Example 3-8</u>: Our inverse pair is:

PSL(b) maximize $\{a = 4x_1+10x_2\}$

$$\text{subject to } 2x_1+5x_2 \leqq b_1$$
$$x_1+x_2 \leqq b_2$$
$$x_1 \leqq b_3$$
$$x_1,x_2 \geqq 0$$

IVL(a) find $\text{Pmax}\begin{Bmatrix} b_1 = 2x_1+5x_2 \\ b_2 = x_1+x_2 \\ b_3 = x_1 \end{Bmatrix}$

$$\text{subject to } 4x_1+10x_2 \geqq a$$
$$x_1,x_2 \geqq 0.$$

We want to use algorithm 3-7 to obtain an optimal inverse triple with $\bar{a} = 20$ and $\bar{b}$ proportional to $(5,1.5,1) = \hat{b}$. So first we solve PSL(5,1.5,1). One optimal solution for this problem is $\hat{x} = (5/6,2/3)$ with $\hat{a} = 10$. Therefore $\gamma = \bar{a}/\hat{a} = 2$, so $\bar{x} = (5/3,4/3)$ solves PSL(10,3,2). But $D\bar{x} = (10,3,5/3)$ so we set $\bar{b} = (10,3,5/3)$ and check to see if $T = (20,(10,3,5/3),(5/3,4/3))$ is an optimal inverse triple. The dual of PSL(10,3,5/3) is

$$\text{minimize } 10y_1+3y_2+5/3y_3$$
$$\text{such that } 2y_1+y_2+y_3 \geqq 4$$
$$5y_1+y_2 \geqq 10$$
$$y_1,y_2,y_3 \geqq 0.$$

This has the unique optimal solution $y = (2,0,0)$ so T does not satisfy the requirements. However PSL(10,3,2) or

(PSL(10,3,5/3)) has another optimal solution $\tilde{x} = (0,2)$ and $D\tilde{x} = (10,2,0)$. So we reset $\tilde{b} = (10,2,0)$ and consider the dual of PSL(10,20,0):

$$\text{minimize } 10y_1+2y_2+0y_3$$
$$\text{such that } 2y_1+y_2+y_3 \geqq 4$$
$$5y_1+y_2 \geqq 10$$
$$y_1, y_2, y_3 \geqq 0.$$

This has a positive optimal solution $(1,5,\alpha)$ for any positive α so the inverse triple $(20,(10,2,0),(0,2))$ is optimal. Since there is no positive basic solution to the dual, the point $x = (0,2)$ is the only efficient solution to IVL(20).

In this case the single triple found is also the entire set of optimal inverse triples. Also $c = (4,10)$ can not be written as a positive combination of the rows of D which are $(2,5)$, $(1,1)$ and $(1,0)$. If we add the row $(-1,0)$ from $-I$ then $1(2,5) + 5(1,1) + \alpha(1,0) + (3+\alpha)(-1,0) = (4,10)$. Therefore in the optimal inverse triples $x_1 = 0$.■

<u>Example 3-9</u>: This is a special case of example 1-8. We are trying to decide what output to produce and how much resources to use to make 5 units of profit under the following conditions: there are 3 resources and 3 outputs, the profit function is $a = cx = x_1+x_2+x_3$ and the technology

matrix is $\begin{pmatrix} 2 & 8 & 3 \\ 1 & 2 & 4 \\ 6 & -12 & 3 \end{pmatrix}$.

Thus we want an optimal inverse triple with $\bar{a} = 5$ for the pair:

PSL(b) maximize $\{a = x_1+x_2+x_3\}$

$$\text{such that } \begin{aligned} 2x_1+8x_2+3x_3 &\leqq b_1 \\ x_1+2x_2+4x_3 &\leqq b_2 \\ 6x_1-12x_2+3x_3 &\leqq b_3 \\ x_1,x_2,x_3 &\geqq 0; \end{aligned}$$

IVL(a) find Pmin $\begin{Bmatrix} b_1 = 2x_1+8x_2+3x_3 \\ b_2 = x_1+2x_2+4x_3 \\ b_3 = 6x_1-12x_2+3x_3 \end{Bmatrix}$

$$\text{subject to } \begin{aligned} x_1+x_2+x_3 &\geqq a \\ x_1,x_2,x_3 &\geqq 0. \end{aligned}$$

It is obvious that the efficient set for IVL(5) is a subset of $S = \{(x_1,x_2,x_3)\,|\,x_1+x_2+x_3 = 5,\ x_1,x_2,x_3 \geqq 0\}$ but it is not obvious which one. We decide that we would like to use the resources in the ratio $\hat{b} = (3,2,1)$ so we consider PSL(3,2,1). An optimal solution for this problem is $\hat{x} = (1/3,1/6,1/3)$ with value $\hat{a} = 5/6$. Thus $\gamma = \frac{5}{5/6} = 6$ so $\bar{x} = (2,1,2)$ solves PSL(18,12,6) with value $\bar{a} = 5$. Also $D\bar{x} = (18,12,6)$ so $T = (5,(18,12,6),(2,1,2))$ is a candidate for the required optimal inverse triple. To check, we consider the dual of PSL(18,12,6):

$$\text{minimize} \quad 18y_1+12y_2+6y_3$$
$$2y_1+y_2+6y_3 \geqq 1$$
$$8y_1+2y_2-12y_3 \geqq 1$$
$$3y_1+4y_2+3y_3 \geqq 1$$
$$y_1, y_2, y_3 \geqq 0$$

This has the unique solution (1/4,0,1/12) so T does not satisfy the requirements. Checking PSL(18,12,6) we find the alternate optimal solution $\tilde{x} = (11/3,4/3,0)$. $D\tilde{x} = (18,19/3,6)$ so we reset $\bar{b}$ to this value and consider the dual of PSL(18,19/3,6):

$$\text{minimize} \quad 18y_1+19/3y_2+6y_3$$
$$\text{such that} \quad 2y_1+y_2+6y_3 \geqq 1$$
$$8y_1+2y_2-12y_3 \geqq 1$$
$$3y_1+4y_2+3y_3 \geqq 1$$
$$y_1, y_2, y_3 \geqq 0$$

This dual has two optimal basic solutions $\hat{y} = (1/4,0,1/12)$ and $\tilde{y} = (0,3/4,1/24)$. We can take y to be any convex combination of $\hat{y}$ and $\tilde{y}$, for example $\bar{y} = (1/8,3/8,1/16)$. Therefore $\bar{b} = (18,19/3,6)$ is the required resource vector. IVL$(5,\bar{y})$ now becomes:

$$\bar{y}\bar{b} = \min \begin{cases} 1/8(2x_1+8x_2+3x_3) \\ +3/8(x_1+2x_2+4x_3) \\ +1/16(6x_1-12x_2+3x_3) \end{cases}$$

$$= x_1+x_2+33/16x_3$$

$$\text{subject to} \quad x_1+x_2+x_3 \geq 5$$
$$x_1, x_2, x_3 \geq 0$$

The set of optimal solutions to this is $E = \{\alpha(5,0,0)+(1-\alpha)(0,5,0) \;\; 0 \leq \alpha \leq 1\}$ which is also the set of efficient solutions to IVL(5). From this we can derive the set of efficient values:

$$B(5) = \{\alpha(10,5,30)+(1-\alpha)(40,10,-60) \;\; 0 \leq \alpha \leq 1\}.$$

The $\bar{b}$ found above corresponds to $\alpha = 11/15$ and the -60 indicates that we can run the processes in such a way as to output b_3 instead of consuming it.

Notice that in this case $\bar{y}$ is not unique and E is a proper subset of S (one edge). Again $c = (1,1,1)$ cannot be written as a positive combination of the rows of D, $(2,8,3)$, $(1,2,4)$ and $(6,-12,3)$. However $c = 1/8(2,8,3) + 3/8(1,2,4) + 1/16(6,-12,3) + 19/16(0,0,-1)$, so we must include the last row of $-I$ and $x_3 = 0$ in the optimal inverse triples.∎

3.2 Vector-Vector Linear Inverse Problems

We now consider what happens when both $P(\bar{b})$ and $I(\bar{a})$ are vector-valued though still linear. For this discussion we define the following pair of problems.

<u>Definition 3-10</u>: The <u>primal problem</u> is:

PVL(b) find $\bar{a} = \text{Pmax}\{a = Cx\}$

such that $Dx \leqq b$

$x \geqq 0.$

The <u>inverse problem</u> is

IVL(a) find $\bar{b} = \{\text{Pmin } b = Dx\}$

such that $Cx \geqq a$

$x \geqq 0.$

In both problems $D \in R^{m\times n}$; $C \in R^{k\times n}$; $b \in R^{m}$; $a \in R^{k}$; $x \in R^{n}$. PVL(b,μ) and IVL(a,λ) are the <u>scalar-valued</u> versions of PVL(b) and IVL(a) respectively. We assume PVL(b) and IVL(a) satisfy the conditions of definition 1-3.∎

We now have two multiplier vectors, μ and λ, to consider. Previously we related λ to the solutions of the dual of PSL(b). Now PVL(b) is vector-valued, and to get a usual dual we have to consider PVL(b,μ) for some $\mu > 0$. We can then get a result similar to that obtained before and a similar result relating μ to solutions of IVL(a,λ). However since the dual of PVL(b,μ) depends on μ and the dual of IVL(a,λ) depends on λ, the multipliers for the two problems are not independent. The following theorem gives us the relationships discussed.

<u>Theorem 3-11</u>: $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple for the inverse pair PVL$(\bar{b})$ and IVL$(\bar{a})$ if and only if there exist $\bar{\mu} > 0$ and $\bar{\lambda} > 0$ such that

(1) $\bar{x}$ solves PVL$(\bar{b},\bar{\mu})$ with optimal value $\bar{\mu}\bar{a}$,

(2) $\bar{x}$ solves IVL$(\bar{a},\bar{\lambda})$ with optimal value $\bar{\lambda}\bar{b}$,

(3) $\bar{\lambda}$ solves the dual of $PVL(\bar{b},\bar{\mu})$,

(4) $\bar{\mu}$ solves the dual of $IVL(\bar{a},\bar{\lambda})$.

Proof:

■ The same theorem for the nonlinear case is theorem 4-15. We prove that theorem in the next chapter using a modified Lagrange function. Since the present theorem is just a special case of the nonlinear result, we delay its proof until chapter 4 where it is a corollary to theorem 4-15.■

We next turn to the problem of finding optimal inverse triples for pairs of vector-valued problems. The next theorem gives the results which enable us to construct algorithms for finding these triples.

Theorem 3-12: Let PVL(b) and IVL(a) be an inverse pair of problems, and let $\hat{b}$ and $\hat{a}$ be elements of $\bar{B}$ and $\bar{A}$ respectively.

(1) If $PVL(\hat{b},\bar{\mu})$ has a finite solution for $\bar{\mu} > 0$, then there is a $\bar{b} \leq \hat{b}$ with $\bar{b} \not< \hat{b}$, a $\bar{\lambda} > 0$ and an $\bar{x} \geq 0$ such that $\bar{x}$ solves $PVL(\bar{b},\bar{\mu})$, $\bar{\lambda}$ solves the dual of $PVL(\bar{b},\bar{\mu})$, $(\bar{a} = C\bar{x},\bar{b},\bar{x})$ is an optimal inverse triple and, for this triple, $\bar{\mu}$ and $\bar{\lambda}$ are the multipliers required by theorem 3-11.

(2) If $IVL(\hat{a},\tilde{\lambda})$ has a finite solution for $\tilde{\lambda} > 0$, then there is an $\tilde{a} \geq \hat{a}$ with $\tilde{a} \not> \hat{a}$, a $\tilde{\mu} > 0$ and an $\tilde{x} \geq 0$ such that $\tilde{x}$ solves $IVL(\tilde{a},\tilde{\lambda})$, $\tilde{\mu}$

solves the dual of $IVL(\tilde{a},\tilde{\lambda})$, $(\tilde{a},\tilde{b} = D\tilde{x},\tilde{x})$ is an optimal inverse triple and, for this triple, $\tilde{\mu}$ and $\tilde{\lambda}$ are the multipliers required by theorem 3-11.

(3) We can find all inverse triples by using (1) or (2) and some $\hat{b}$ or $\hat{a}$.

<u>Proof</u>:

(1) Consider the set of all optimal solutions of $PVL(\hat{b},\bar{\mu})$ and take the Pareto minima over x in this set of $b = Dx$. At least one Pareto minimum is at an extreme point. Let this point be $\bar{b}$ and the x that yields it be $\bar{x}$. Now $\bar{x}$ solves $PVL(\bar{b},\bar{\mu})$ and there is no other solution x with $Dx \leq \bar{b}$ since this x would also solve $PVL(\hat{b},\bar{\mu})$. Now $\bar{b} \leq \hat{b}$ and since definition 1-3 is satisfied $\bar{b} \not> \hat{b}$. For $PVL(\bar{b},\bar{\mu})$ and its dual there are a pair of solutions (x,y), by strict complementary slackness, such that $y_i > 0$ if and only if $(Dx)_i = \bar{b}_i$ and $x_j > 0$ if and only if $(yD)_j = (\bar{\mu}C)_j$. But we have seen that for all solutions of $PVL(\bar{b},\bar{\mu})$, $(Dx)_i = \bar{b}_i$ for all $i = 1,\ldots,m$. Therefore there is a $y > 0$ that solves the dual, let it be $\bar{\lambda}$.

We now show $\bar{x}$ solves $IVL(\bar{a},\bar{\lambda})$ where $\bar{a} = C\bar{x}$ and $\bar{\mu}$ solves its dual. $IVL(\bar{a},\bar{\lambda})$ is:

$$\text{minimize } \{\bar{\lambda}b = \bar{\lambda}Dx\}$$
$$\text{subject to } Cx \geq \bar{a}$$
$$x \geq 0.$$

$\bar{x}$ is certainly feasible so if $\bar{x}$ is not optimal then there is an x such that:

$$\bar{\lambda}Dx < \bar{\lambda}\bar{b}$$
$$Cx \geq \bar{a}$$
$$x \geq 0.$$

The value of $PVL(\bar{b},\bar{\mu})$ is $\bar{\mu}C\bar{x} = \bar{\mu}\bar{a}$. The value of its dual is $\bar{\lambda}\bar{b}$. Therefore $\bar{\mu}\bar{a} = \bar{\lambda}\bar{b}$ so there must be an x that solves

$$\bar{\lambda}Dx < \bar{\mu}\bar{a}$$
$$Cx \geq \bar{a}$$
$$x \geq 0$$

Since $\bar{\mu} > 0$, $\bar{\mu}Cx \geq \bar{\mu}\bar{a}$ so

$$\bar{\lambda}Dx < \bar{\mu}Cx.$$

However the constraints on the dual of $PVL(\bar{b},\bar{\mu})$ are $yD \geq \bar{\mu}C$. $\bar{\lambda}$ is a solution to that dual so $\bar{\lambda}D \geq \bar{\mu}C$ and $\bar{\lambda}Dx \geq \bar{\mu}Cx$ for all $x \geq 0$. Thus we have a contradiction and so $\bar{x}$ solves $IVL(\bar{a},\bar{\lambda})$. The dual of $IVL(\bar{a},\bar{\lambda})$ is:

$$\text{maximize } z\bar{a}$$
$$zC \leq \bar{\lambda}D$$
$$z \geq 0.$$

$\bar{\mu}$ is feasible in this problem and $\bar{\mu}\bar{a} = \bar{\lambda}\bar{b} = \bar{\lambda}D\bar{x}$ which is the value of $IVL(\bar{a},\bar{\lambda})$ so $\bar{\mu}$ is an optimal solution.

Therefore $\bar{a},\bar{b},\bar{x},\bar{\lambda},\bar{\mu}$ satisfy all four conditions of theorem 3-11 and $(\bar{a},\bar{b},\bar{x})$ must be an optimal inverse triple with $\bar{\mu}$ and $\bar{\lambda}$ the required multipliers.

(2) similarly.

(3) Assume $(\hat{a},\hat{b},\hat{x})$ is an optimal inverse triple. Then there exist $\hat{\lambda} > 0$, $\hat{\mu} > 0$ by theorem 3-11 such that

conditions (1) to (4) in that theorem are met. Now follow part (1). Since $PVL(\hat{b},\hat{\mu})$ does have a finite solution we need only show that $\bar{b} = \hat{b}$. We can do this by showing that there is no $\bar{x}$ optimal in $PVL(\hat{b},\hat{\mu})$ such that $D\bar{x} \leq b$. Since $\hat{\lambda} > 0$ we would then have $\hat{\lambda}Dx < \hat{\lambda}\hat{b}$. But, as in part (1), $\hat{\mu}\hat{a} = \hat{\lambda}\hat{b}$ and, since $\bar{x}$ is optimal $\hat{\mu}\hat{a} = \hat{\mu}C\bar{x}$, so

$$\hat{\lambda}D\bar{x} < \hat{\mu}\hat{a} = \hat{\mu}C\bar{x}.$$

But $\hat{\lambda}$ solves the dual of $PVL(\hat{b},\hat{\mu})$ so $\hat{\lambda}D \geq \hat{\mu}C$. Therefore there is no $x \geq 0$ that gives $\hat{\lambda}Dx < \hat{\mu}Cx$ so $\bar{x}$ cannot exist. Hence $(\hat{a},\hat{b},\hat{x})$ can be found using $\hat{b}$ and $\hat{\mu}$ in part (1) (or $\hat{a}$ and $\hat{\lambda}$ in part (2)).■

This theorem allows us to develop the following algorithm.

<u>Algorithm 3-13</u>: This algorithm finds an optimal inverse triple for the inverse pair of problems $PVL(b)$ and $IVL(a)$ starting with an estimated value for b labelled $\hat{b}$ and a specified multiplier $\hat{\mu}$. It further determines the set of all multipliers which yield the same triple.

(1) Choose $\bar{\mu} > 0$, $\hat{b} \in B$, set $\bar{b} = \hat{b}$ and solve the linear programming problem $PVL(\bar{b},\bar{\mu})$ for $\bar{x}$. If no finite solution exists, reset $\bar{\mu}$.

(2) Set $\bar{a} = C\bar{x}$ and, if $D\bar{x} \leq \bar{b}$, reset $\bar{b} = D\bar{x}$.

(3) Solve the linear programming problem which is the dual of $PVL(\bar{b},\bar{\mu})$. If it has a positive solution $\bar{\lambda}$, $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. If not, there is another extreme point, x, of the set of optimal solutions of $PVL(\bar{b},\bar{\mu})$ with $Dx \leq \bar{b}$.

Reset $x = \bar{x}$ and return to (2).

(4) Let $M = \{\mu/||\mu||\}$. Let e_i be the standard unit vector and, if $e_i D\bar{x} = \hat{b}_i$, solve $IVL(\bar{a},e_i)$. If $\bar{x}$ solves $IVL(\bar{a},e_i)$, solve its dual. If y is an extreme point optimal solution of the dual add $y/||y||$ to M. If $\bar{x}$ does not solve $IVL(\bar{a},e_i)$, some single objective $e_j Cx$ of $PVL(\hat{b})$ is optimized at $\bar{x}$. Put all such e_j in M. Let MU be the convex hull of M. Now the set of all $\mu \in MU$, $\mu > 0$ are all the normalized multiplier vectors that, with $\hat{b}$, lead to the same optimal inverse triple $(\bar{a},\bar{b},\bar{x})$.

(5) Another optimal inverse triple can be found by repeating the algorithm with $\mu > 0$, $\mu \notin MU$.■

Steps (1), (2) and (3) are similar to the procedures used for the scalar-vector case in section 3.1. In step (4) we want to determine those $\mu > 0$ such that the hyperplane $\mu Cx = \mu\bar{a}$ touches the feasible set of $PVL(\hat{b})$ at $\bar{x}$ but does not intersect the interior of the efficient set.

$IVL(\bar{a},e_i)$ is:

$$\begin{aligned} &\text{minimize } \{e_i Dx = b_i\} \\ &\text{subject to } \quad Cx \geq \bar{a} \\ &\qquad\qquad\qquad x \geq 0, \end{aligned}$$

its dual is:

$$\text{maximize } y\bar{a}$$
$$\text{subject to } yC \leq e_iD$$
$$y \geq 0.$$

Assume $\bar{x}$ solves $IVL(\bar{a},e_i)$ and $\bar{y}$ solves its dual. So $\bar{y}\bar{a} = e_iD\bar{x} = \hat{b}_i$. Consider $PVL(\hat{b},\bar{y})$ which is:

$$\text{maximize } \{\bar{y}Cx = \bar{y}a\}$$
$$\text{subject to } Dx \leq \hat{b}$$
$$x \geq 0,$$

and its dual which is:

$$\text{minimize } z\hat{b}$$
$$\text{subject to } zD \geq \bar{y}C$$
$$z \geq 0.$$

$\bar{x}$ is feasible in $PVL(\hat{b},\bar{y})$, e_i is feasible in its dual. $\bar{y}C\bar{x} = \bar{y}\bar{a} = e_iD\bar{x} = \hat{b}_i = e_i\hat{b}$ so these solutions are optimal. Therefore if $\bar{x}$ solves $IVL(\bar{a},e_i)$, $\bar{x}$ is optimal in $PVL(\hat{b},\bar{y})$. So we get the same optimal inverse triple. If $\bar{x}$ is not a solution of $IVL(\bar{a},e_i)$, e_iD is not a positive linear combination of the rows of C. If it was we would have $e_iD = \hat{y}C$ for some $\hat{y} > 0$. $\bar{x}$ is feasible in $IVL(\bar{a},e_i)$ and $\hat{y}$ is feasible in its dual. But $\hat{y}C\bar{x} = \hat{y}\bar{a} = e_iD\bar{x} = \hat{b}_i$ so $\bar{x}$ and $\hat{y}$ are actually optimal. This would contradict $\bar{x}$ not being a solution. If e_iD is not a positive linear combination of the rows of C, one of the objective functions $e_j(Cx = a)$ must be optimized at $\bar{x}$ since $\bar{x}$ is efficient in $PVL(\hat{b})$. If $\bar{x}$ is optimal in $PVL(\hat{b},\hat{\mu})$ and $PVL(\hat{b},\tilde{\mu})$, it is optimal in $PVL(\hat{b},\alpha\hat{\mu}+(1-\alpha)\tilde{\mu})$ where $0 \leq \alpha \leq 1$. Hence $\bar{x}$ is optimal in $PVL(\hat{b},u)$ for all μ in MU. If $\mu \in MU$ and $\mu > 0$, a $\lambda > 0$ which solves the dual of PVL($\bar{b},\mu$) and is the other required multiplier vector must exist by theorem 3-12. Hence the positive elements of MU are the other multipliers that lead to the same optimal inverse triple. Each triple found by this algorithm has an $\bar{x}$ that is an extreme point of the feasible set of PVL($\hat{b}$). If $\bar{x}$ and $\hat{x}$ are efficient extreme points, and if $\alpha\bar{x} + (1-\alpha)\hat{x}$ is on the boundary of the feasible set, it is also an efficient point. In such a case, if $a = C(\alpha\bar{x} + (1-\alpha)\hat{x})$ and $b = D(\alpha\bar{x} + (1-\alpha)\hat{x})$, $(a, b, \alpha\bar{x} + (1-\alpha)\hat{x})$ is also an optimal inverse triple. Using The algorithm and this argument, we can find all optimal inverse triples that can be obtained starting with $\hat{b}$. We can, of course, develop an equivalent algorithm starting with an $\hat{a} \in \bar{A}$ and a $\bar{\lambda} > 0$.

The following example demonatrates the method discussed above.

<u>Example 3-14</u>: Our primal problem is

PVL(b)
$$\bar{a} = \text{Pmax}\begin{Bmatrix} a_1 = 2x_1+x_2 \\ a_2 = 2x_1+3x_2 \end{Bmatrix}$$
$$\text{subject to } \begin{aligned} x_1+x_2 &\leq b_1 \\ 2x_1+5x_2 &\leq b_2 \\ 6x_1+x_2 &\leq b_3 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Our inverse problem is

IVL(a). $$\bar{b} = \mathrm{Pmin}\begin{Bmatrix} b_1 = x_1+x_2 \\ b_2 = 2x_1+5x_2 \\ b_3 = 6x_1+x_2 \end{Bmatrix}$$

$$\text{subject to } \quad 2x_1+x_2 \geqq a_1$$
$$2x_1+3x_2 \geqq a_2$$
$$x_1, x_2 \geqq 0.$$

We decide we want $\bar{b}$ to be as close as possible to $b = (7,30,24)$. Figure 3-3 shows the constraint set for PVL(7,30,24). Since a_1 is optimized at $(17/5,18/5)$ and a_2 is optimized at $(5/3,16/3)$, the efficient set must be $\{(x_1x_2) \mid x_1+x_2 = 7; 5/3 \leqq x_1 \leqq 17/5, 18/5 \leqq x_2 \leqq 16/3\}$. Let us first pick $\bar{\mu} = (3/4,1/4)$. This leads to $\bar{\mu}\bar{a} = 61/5$ and $\bar{x} = (17/5,18/5)$ so $\bar{a} = (52/5,88/5)$ and $D\bar{x} = (7,124/5,24)$. Our candidate for an optimal triple is then $T_1 = ((52/5,88/5),(7,124/5,24),(17/5,18/5))$. To test it we check the dual of PVL((7,124/5,24),(3/4,1/4))

$$\min \quad 7y_1+124/5y_2+24y_3$$
$$y_1+2y_2+6y_3 \geqq 2$$
$$y_1+5y_2+y_3 \geqq 1.5$$
$$y_1,y_2,y_3 \geqq 0.$$

This has optimal basic solutions $(0,1/4,1/4)$ and $(7/5,0,1/10)$ so it has a positive optimal solution (e.g. $(7/10,1/8,7/40)$) and T_1 satisfies our requirements. To

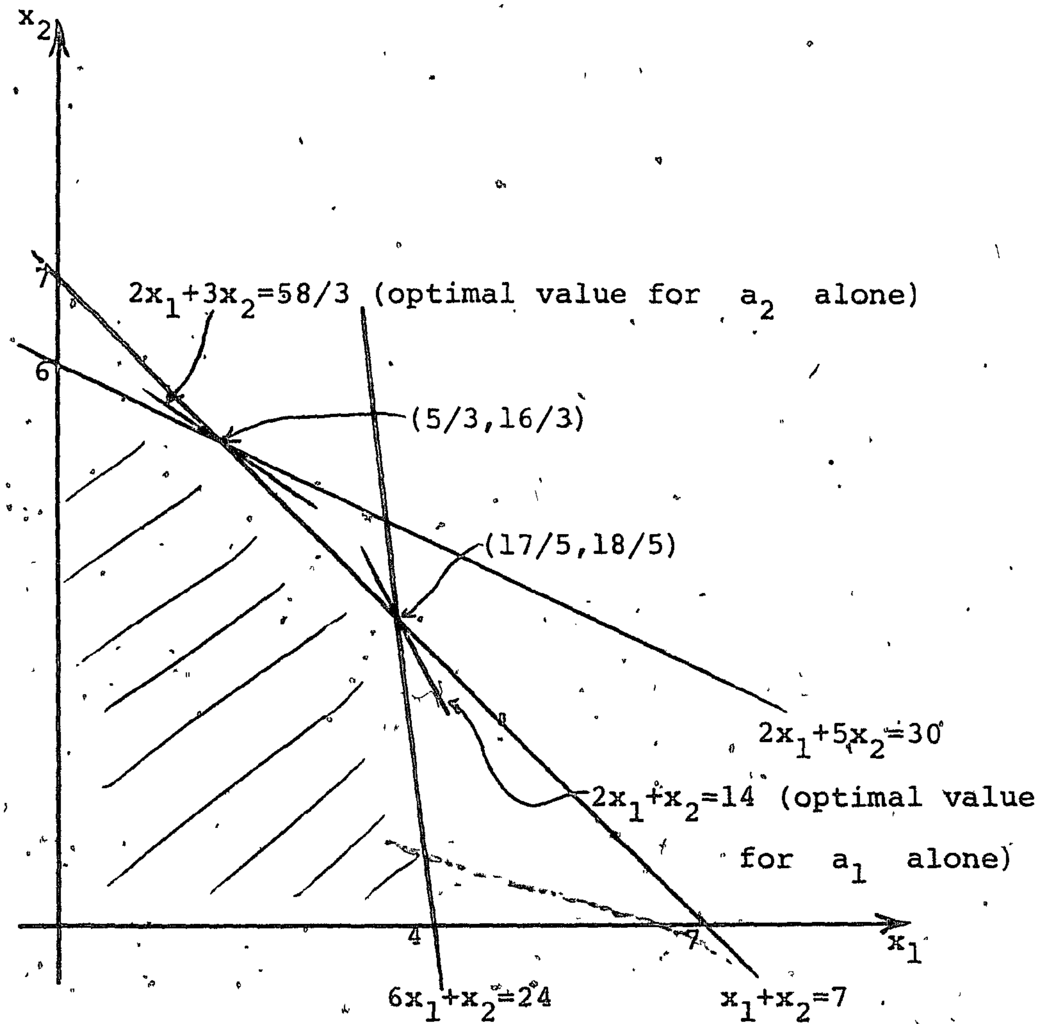


Figure 3-3: Constraint set for PVL(7,30,24) in example 3-14 showing optimal solutions if each objective is taken separately.

find what other μ's yield the same triple, since the first and third constraints in the original problem are tight we study the two problems:

$$\text{IVL}((52/5,88/5),(1,0,0)) \quad \text{minimize } x_1+x_2$$
$$\text{such that } 2x_1+x_2 \geq 52/5$$
$$2x_1+3x_2 \geq 88/5$$
$$x_1,x_2 \geq 0$$

$$\text{IVL}((52/5,88/5),(0,0,1)) \quad \text{minimize } 6x_1+x_2$$
$$\text{such that } 2x_1+x_2 \geq 52/5$$
$$2x_1+3x_2 \geq 88/5$$
$$x_1,x_2 \geq 0$$

$\bar{x}$ is optimal in the first but not the second. The solution to the dual of the first is $\bar{z} = (1/4,1/4)$. (Notice that $e_1D = (1/4,1/4)C$.) e_3D cannot be written as a positive linear combination of the rows of C and hence some $e_jCx = a_j$ is optimized at $\bar{x}$. We already know (see figure 3-3) that the first objective is optimized at this point. Therefore, for the triple we have found, $MU = \{(\mu_1,\mu_2) = \alpha(1/2,1/2) + (1-\alpha)(0,1) \mid 0 < \alpha \leq 1\}$. (We exclude the single point $(0,1)$ since μ must be positive.) We next pick a μ outside this set - say $\bar{\mu} = (3/4,1/4)$ and resolve. Following the above procedure we get a new triple $T_2 = ((26/3,58/3),(7,30,46/3),(5/3,16/3))$. And this is optimal for the set of normalized μ's $\{(\mu_1,\mu_2) = \alpha(1/2,1/2) + (1-\alpha)(1,0) \mid 0 < \alpha \leq 1\}$. These two sets of μ's cover all possible values so T_1 and T_2 are the extreme points to the set of optimal inverse triples. Therefore the complete set of optimal triples is

$$\{(a,b,x) \mid a = \alpha(52/5,88/5) + (1-\alpha)(26/3,58/3),$$
$$b = \alpha(7,124/5,24) + (1-\alpha)(7,30,46/3),$$
$$x = \alpha(17/5,18/5) + (1-\alpha)(5/3,16/3) \mid 0 \le \alpha \le 1\}$$

Since $\mu = (1/2,1/2)$ is in both sets, the other efficient solutions to PVL(7,30,14) can be found from solving PVL((7,30,24),(1/2,1/2)) and from these the other optimal triples can be derived.∎

If we don't have a particular desired $\hat{b}$ (or $\hat{a}$) we may have to try an iterative approach to the problem. We can pick one or the other and proceed as above. Because of linearity, if (a,b,x) is an optimal inverse triple, so is $\alpha(a,b,x)$ for all $\alpha > 0$. So once finished we can pick an a or b that is not a positive multiple of any of those in the optimal triples found, and repeat. In this way we can get an idea of what happens as we alter the ratio of the components in a or b. Once this is done we can pick a triple from amongst those derived.

We discuss one last point in this section. It would be nice if the following result held when $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple:

(1) Given $\hat{a}, \tilde{a} \in A^I(\bar{b})$ then $B^I(\hat{a}) = B^I(\tilde{a})$,

(2) Given $\hat{b}, \tilde{b} \in B^I(\bar{a})$ then $A^I(\hat{b}) = A^I(\tilde{b})$.

However this does not hold in general. What does happen instead is demonstrated in the following example.

Example 3-15: Our primal problem is:

PVL(b)

$$\bar{a} = \text{Pmax}\left\{\begin{array}{l} a_1 = 2x_1+2x_2+x_3 \\ a_1 = 2x_1+x_2+2x_3 \end{array}\right\}$$

$$\text{subject to} \quad \begin{array}{l} x_1+2x_2+2x_3 \leq b_1 \\ 2x_1+x_2+x_3 \leq b_2 \\ x_1,x_2,x_3 \geq 0 \end{array}$$

Starting with $\bar{b} = (2,2)$, we get the set of optimal inverse triples

$$T = \{(\alpha(8/3,2)+(1-\alpha)(2,8/3)),(2,2),\alpha(2/3,2/3,0) + (1-\alpha)(2/3,0,2/3))\,|\,0 \leq \alpha \leq 1\}$$

so $A^I(2,2) = \{\alpha(8/3,2)+(1-\alpha)(2,8/3)\,|\,0 \leq \alpha \leq 1\}$

Picking $\bar{a} = (8/3,2) \in A^I(2,2)$ and studying the inverse problem:

IVL(8/3,2)

$$\bar{b} = \text{Pmax}\left\{\begin{array}{l} b_1 = x_1+2x_2+2x_3 \\ b_2 = 2x_1+x_2+x_3 \end{array}\right\}$$

$$\text{subject to} \quad \begin{array}{l} 2x_1+2x_2+x_3 \geq 8/3 \\ 2x_1+x_2+2x_3 \geq 2 \\ x_1,x_2,x_3 \geq 0 \end{array}$$

we obtain $B^I(8/3,2) = \{\alpha(2,2)+(1-\alpha)(28/9,14/9)\,|\,0 \leq \alpha \leq 1\}$.

Now take $\hat{b} = (28/9,14/9)$ and solve PVL(28/9,14/9).

This yields

$A^I(28/9,14/9) = \{\alpha(28/9,14/9)+(1-\alpha)(14/9,28/9)\,|\,0 \leq \alpha \leq 1\}$.

$(2,2)$ and $(28/9,14/9)$ are both in $B^I(8/3,2)$ but

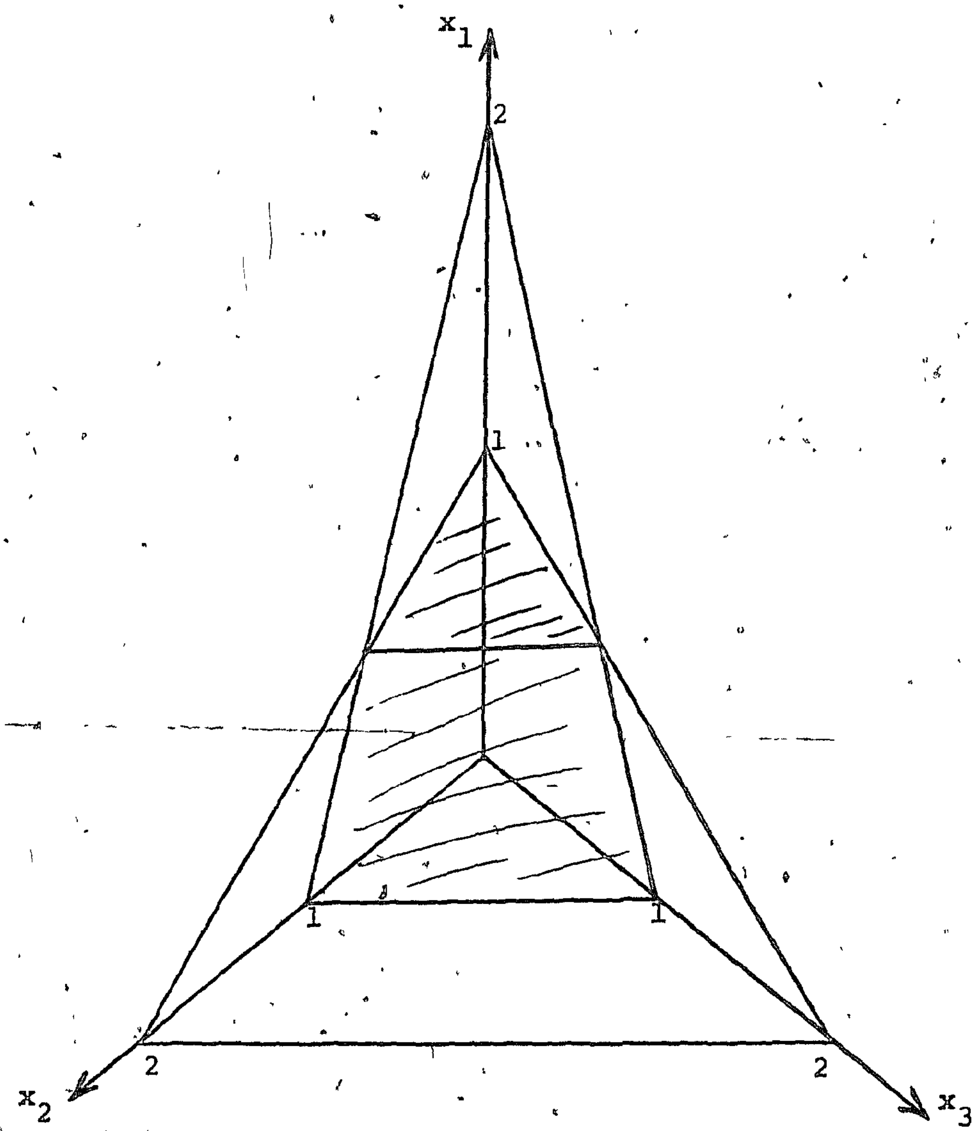


Figure 3-4: Surface of the constraint set for PVL(2,2).

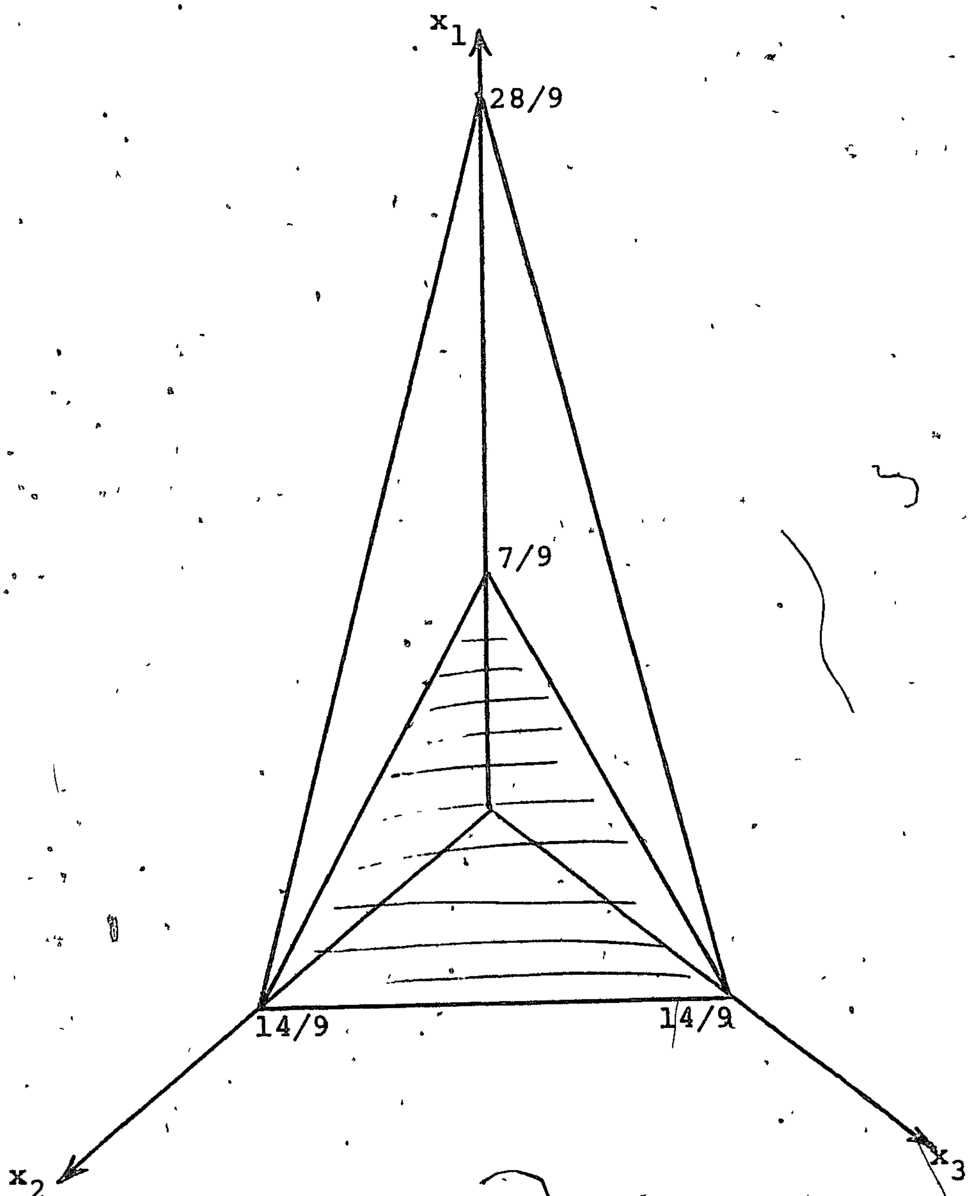


Figure 3-5: The constraint set for PVL(28/9,14/9).

$A^I(2,2) \neq A^I(28/9,14/9)$. However notice that $A^I(2,2) \subset A^I(28/9,14/9)$ (e.g. $(8/3,2) = 5/7(28/9,14/9)+2/7(14/9,28/9)$ and $(2,8/3) = 2/7(28/9,14/9)+5/7(14/9,28/9)$). Compare figures 3-4 and 3-5. In both, the set of efficient points is the segment of the line of intersection between the two planes cut off by the positive orthant and the sets $A^I(2,2)$ and $A^I(28/9,14/9)$ are the values found from x's on these segments. As the ratio of components of b is shifted the intersection line shifts and the set $A^I(b)$ expands or shrinks as more or less of the line is in the positive orthant.∎

Though $A^I(b_1)$ and $A^I(b_2)$ may not be equal the example suggests the following. Find $A^I(b)$ for some b. Determine $B^I(a)$ for all $a \in A^I(b)$. Continue working back and forth finding $B^I(a)$ for all a's in all $A^I(b)$ found and finding $A^I(b)$ for all b's in all $B^I(a)$ found. Let A be the union of all $A^I(b)$ and let B be the union of all $B^I(a)$. For all $a \in A$, $B^I(a) \subset B$ and for all $b \in B$, $A^I(b) \subset A$. If $\bar{a} \in A$, there is no $a \in A$, $a \geq \bar{a}$. If $\bar{b} \in B$, there is no $b \in B$, $b \leq \bar{b}$. Since, as in this example, there may be an uncountable number of b's in B and a's in A, we cannot guarantee that we can produce the entire sets A and B. In the above example $A = A^I(28/9,14/9)$ but in general A may not equal any particular $A^I(b)$ (similarly B may not equal any particular $B^I(a)$).

3.3: Efficiency of LP Solutions

This section considers a problem that may sometimes exist with the normal solution concept in linear programming. Instead of considering just the objective function value we include a consideration of the tightness of the constraints. The normal solution technique does not include this and thus leaves out a discussion of the efficiency with which the solution is obtained. Two methods of surmounting the problem are discussed.

Consider the following example.

Example 3-16: maximize $4x_1+10x_2$

$$\begin{aligned} \text{such that}\quad 2x_1+5x_2 &\leq 10 \\ x_1+x_2 &\leq 3 \\ x_1,x_2 &\geq 0 \end{aligned}$$

For ease of exposition let us assume that we want to determine the amount of two products to produce from a given, fixed amount of two resources. The optimal basic solutions for this problem are (5/3,4/3) and (0,2) (see figure 3-6). Both give an optimal objective function value of 20. However for the first solution the constraints are tight while for the second solution there is one surplus unit of the second resource.

The normal linear programming solution to this problem is that the given optimal vectors (0,2) and (5/3,4/3) (and

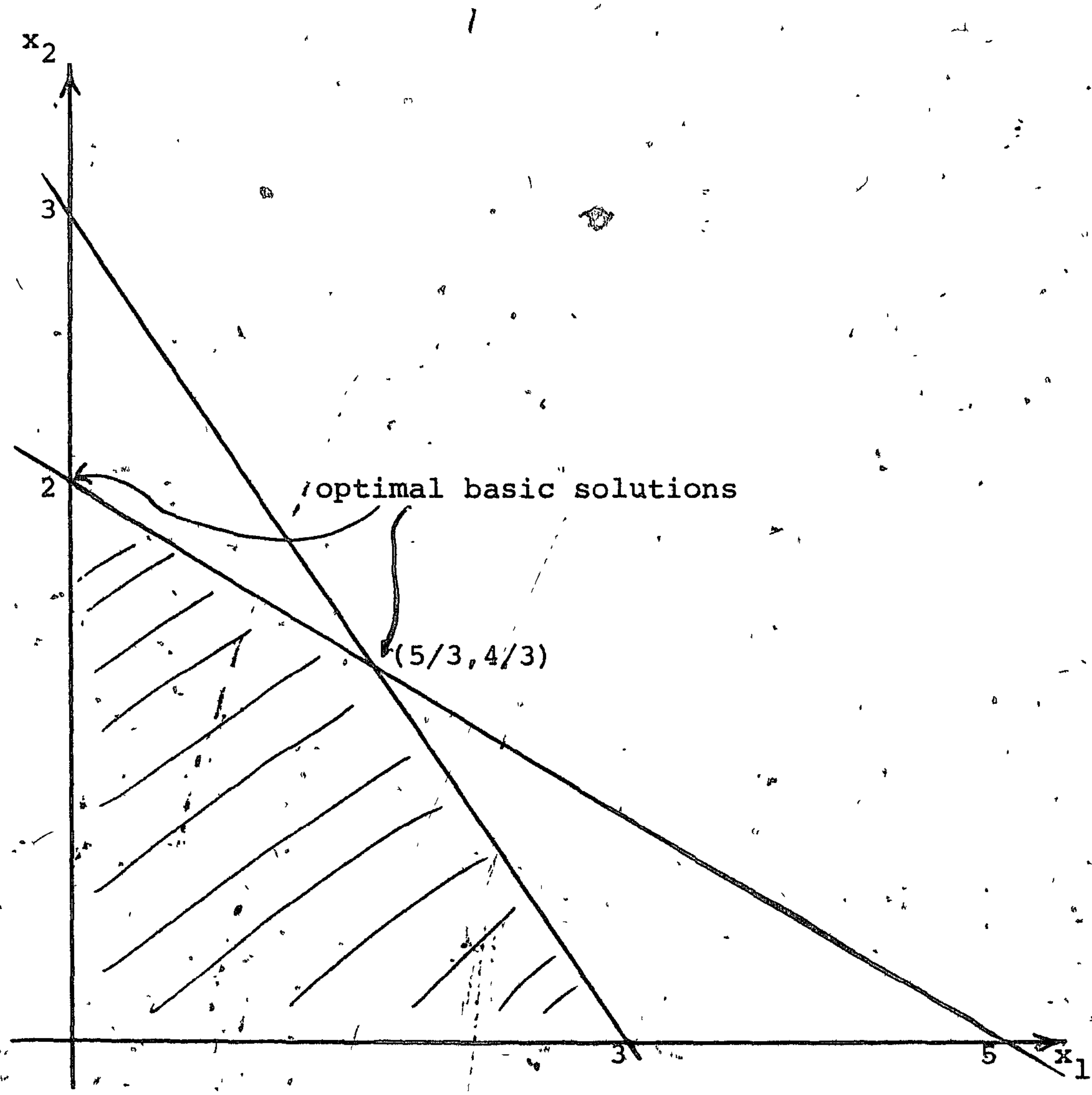


Figure 3-6: Constraint set for example 3-16 showing optimal basic solutions.

any convex combination of them) are equivalent. But accepting this solution implies a number of assumptions that are often ignored. First we assume that the resources are freely available at the level stated except for any costs absorbed into the objective function through x_1 and x_2. Second we assume that a surplus of either resource is costless (there are no storage charges, etc.) and profitless (it cannot be sold or used for other purposes). Third we assume that all costs and revenues are reflected in the objective function. The distribution of any surplus resources may be a secondary consideration, but it seems that (assuming storage costs are not critical) the solution $(0,2)$ is likely to be considered superior to $(5/3,4/3)$: we can obtain the same profit and still have 1 unit of resource 2 left over. That is $(0,2)$ is a more efficient solution.

The dual of this problem is

$$\text{minimize } 10y_1+3y_2$$
$$\text{such that } 2y_1+y_2 \geq 4$$
$$5y_1+y_2 \geq 10$$
$$y_1, y_2 \geq 0$$

This has the unique optimal solution $(2,0)$. This indicates that the second resource has a value of zero in the production of the two outputs. It does not mean that the surplus has no value for alternate uses. From the discussion in section 3.1, notice that though $(5/3,4/3)$ uses up all resources $(D\bar{x} = \bar{b})$ we cannot find a positive set of dual variables. $\bar{x} = (5/3, 4/3)$ is not part of an optimal inverse triple - to get one here we need to cut the second resource to 2 units and the triple is $(a = 20,\ b = (10,2),\ x = (0,2))$.

The simplex method, as usually performed, cannot distinguish between the two basic solutions found. However a minor extension allows us to check both the objective value and the tightness of constraints. First we add the slack variables x_3 and x_4 to the problem. Usually they are given a weight of zero in the objective function. Instead give x_3 weight ε_3 and x_4 weight ε_4. Our problem is now:

$$\text{maximize} \quad 4x_1+10x_2+\varepsilon_3x_3+\varepsilon_4x_4$$

$$\text{such that} \quad 2x_1+5x_2+x_3 = 10$$

$$x_1+x_2+x_4 = 3$$

$$x_1, x_2, x_3, x_4 \geqq 0$$

The value at $(x_1x_2) = (5/3, 4/3)$ is still 20 but at $(0,2)$ it is $20 + \varepsilon_4$. Since nothing in the problem indicates the value of a surplus in one resource relative to the other, we should really consider the objective in vector terms. If V is the optimal value of the original problem then we compare $(V, \alpha_1\varepsilon_3, \beta_1\varepsilon_4)$ and $(V, \alpha_2\varepsilon_3, \beta_2\varepsilon_4)$. In the present problem the point $(0,2)$ with value $(20,0,\varepsilon_4)$ is still better than $(5/3, 4/3)$ with value $(20,0,0)$. However if we had to compare $(20,\varepsilon_3,0)$ and $(20,0,\varepsilon_4)$ we could not rate the points without further information.

Another way to handle the problem is as follows. We noted that $x = (0,2)$ was part of an inverse optimal triple while $x = (5/3,4/3)$ was not. We can pick an optimal $\hat{x}$, set $\hat{b} = D\hat{x}$ and $\hat{a}$ to the optimal value, and use theorem 3-3 to check if $(\hat{a},\hat{b},\hat{x})$ is an optimal inverse triple. If it is, then there is no alternate optimal x that gives at least as much surplus in each resource as x and more in some. If it is not, then we can test some other optimal solution.∎

This example demonstrates a possible method of rating alternate optima for an LP problem. This method is not restricted to $D \in R^{2x2}$ but works for any $D \in R^{mxn}$, m and n finite, as long as the primal problem and its natural inverse (find $\bar{b} = \text{Pmin}\{b = Dx\}$, $cx \geqq a$, $x \geqq 0$) satisfy definition 1-3.

Chapter Four - Nonlinear Inverse Pairs

In this chapter we consider nonlinear inverse pairs of mathematical programming problems. First we discuss several known results for the scalar-valued nonlinear programming problem defined by:

Definition 4-1: The primal scalar-valued problem is

PSN(b) maximize $f(x)$

subject to $g(x) \leq b$

$x \in K \subseteq R^n$, K nonempty and convex,

$f(x): R^n \to R$, $g(x): R^n \to R^m$,

$f(x)$ concave on K,

$g_i(x)$ convex on K $(i = 1, \ldots m)$. ■

For these we need a dual for PSN(b) and a property introduced as "bounded steepness" in [12] and developed in [13] where it is called "stability". Then we apply these results to the following inverse pair of problems in which either problem may be vector-valued.

Definition 4-2: Our primal problem is

PVN(b) $\bar{a} = \text{Pmax}\{a = f(x)\}$

subject to $g(x) \leq b$, $x \in K$.

Our inverse problem is

IVN(a) $\bar{b} = \text{Pmin}\{b = g(x)\}$

subject to $f(x) \geq a$, $x \in K$.

K is a convex nonempty subset of R^n, $g(x): R^n \to R^m$, $f(x): R^n \to R^k$, $f_j(x)$ $(j = 1, \ldots, k)$ is concave on K

and $g_i(x)$ $(i = 1,\ldots,m)$ is convex on K. We assume PVN(b) and IVN(a) satisfy the conditions of definition 1-3. The scalar-valued versions are denoted by PVN(b,μ) and IVN(a,λ) respectively.■

4-1: Review of Nonlinear Programming and Duality

For the nonlinear problem PSN(b) various duals are proposed in the literature (see, for example, [13, 23, 26, 28, 29]. These vary depending on restrictions imposed on the problem and the strength of the results obtained. We use the dual presented by Geoffrion in [13]. A comparison between this dual and others is given in Geoffrion's paper. When restricted to the linear case discussed in chapter 3 all the proposed duals reduce to the usual linear programming dual. Our dual then is defined as follows.

Definition 4-3: The dual of PSN(b) is

DSN(b) minimize $\{\text{supremum } f(x) - u(g(x)-b) \mid x \in K\}$

subject to $u \geq 0$

f, g, K as above, $u \in R^m$.■

Notice that the minimand in DSN(b) is convex in u since it is the pointwise supremum of a collection of functions linear in u.

Given these definitions, consider the following example taken from [13].

Example 4-4: We consider the problem:

PSN(0) maximize $\sqrt{x}$

subject to $g(x) = x \leq 0$

$x \in K = R^+$.

This problem has an optimal value of 0 at $x = 0$, the only feasible point. The dual of this problem is

$$\text{DSN}(0) \qquad \underset{u \geq 0}{\text{minimize}} \{ \sup_{x \geq 0} (\sqrt{x} - u(x-0)) \}$$

$$= \min_{u \geq 0} \begin{cases} \infty & \text{if } u = 0 \\ \frac{1}{4u^2} & \text{if } u > 0 \end{cases}$$

This set has an infimal value of 0 but this is not taken on by any finite $u \geq 0$. Therefore there is no solution to the dual. If we let the right hand side of the constraint be b instead of 0, then PSN(b) is the primal problem of a nonlinear inverse pair with $g(x) = x$, $f(x) = \sqrt{x}$, $K = R^+$, $\mathcal{T} = R^+ \times R^+$, $\mathcal{A} = R^+$, $\mathcal{B} = R^+$. ∎

Generally in nonlinear programming, for a solution to the primal to imply the existence of a solution to the dual we need some kind of qualification on the problem. Usually a constraint qualification is used such as the Kuhn-Tucker Constraint Qualification or the Slater Condition. In [12], Gale proposes an alternative that provides a necessary as well as sufficient condition and thus is the weakest condition obtainable. Geoffrion in [13] calls this "stability" and uses it to provide a new look at duality. First we define a new function.

<u>Definition 4-5</u>: The <u>value</u> <u>function</u> $v(b)$ associated with $PSN(b)$ for all $b \in \mathcal{B}$ is defined as

$$v(b) = \sup\{f(x) \mid g(x) \leq b,\ x \in K\}. \blacksquare$$

The definition of stability is then:

<u>Definition 4-6</u>: $PSN(b)$ is <u>stable</u> at $\bar{b} \in \mathcal{B}$ if $v(\bar{b})$ is finite and there exists an $M > 0$ such that

$$\frac{v(b) - v(\bar{b})}{\| b - \bar{b} \|} \leq M \quad \text{for all } b \in \mathcal{B} - \{\bar{b}\}$$

where $\| \cdot \|$ is any norm on R^m. $\blacksquare$

<u>Example 4-4, cont'd</u>: For this example $v(b) = \sqrt{\bar{b}}$ for all $b \in \mathcal{B}$, so

$$\frac{v(b) - v(0)}{\| b \|} = \frac{\sqrt{\bar{b}} - 0}{b} \qquad b > 0$$

$$= \frac{1}{\sqrt{\bar{b}}} \qquad b > 0.$$

Since $\frac{1}{\sqrt{\bar{b}}} \to \infty$ as $b \to 0$, there is no bound on $\frac{v(b) - v(0)}{\| b \|}$ for all $b \in \mathcal{B}$ and $PSN(0)$ is unstable. $\blacksquare$

Actually, under the assumptions of convexity and concavity on K, g and f, instability of $PSN(\bar{b})$ implies that there is no interior to the set of feasible points. Hence the Slater Condition is sufficient to prove stability.

Notice the parallel between the definition of stability and the definition of proper efficiency. In the former the ratio of changes in the objective function to changes in the constraint bounds must be bounded; in the latter the ratio of changes in one objective to changes in other objectives must be bounded.

With the above tools we can prove the following results taken directly from [13]. If $\bar{b} \in B$, there is an $\bar{x}$ and an $\bar{a}$ such that $f(\bar{x}) \geq \bar{a}$, $g(\bar{x}) \leq \bar{b}$, $\bar{x} \in K$. Hence $g(\bar{x}) \leq \bar{b} + \alpha e$ for all $\alpha > 0$ where e is the vector of 1's of appropriate length and $(\bar{a}, \bar{b} + \alpha e) \in T$ and $\bar{b} + \alpha e \in B$.

Lemma 4-7: $PSN(\bar{b})$ is stable if and only if $v(\bar{b})$ is finite and there exists a scalar $M > 0$ such that

$$\frac{v(\bar{b} + \alpha e) - v(\bar{b})}{\alpha} \leqq M \qquad \text{for all} \quad \alpha > 0. \blacksquare$$

Theorem 4-8: (1) If $\bar{x}$ is feasible in $PSN(\bar{b})$ and $\bar{u}$ is feasible in $DSN(\bar{b})$, then the objective function of $PSN(\bar{b})$ evaluated at $\bar{x}$ is not less than the objective function of $DSN(\bar{b})$ evaluated at $\bar{u}$.

(2) If $PSN(\bar{b})$ is stable, then:

(a) $DSN(\bar{b})$ has an optimal solution,

(b) the optimal values of $PSN(\bar{b})$ and $DSN(\bar{b})$ are equal,

(c) $\bar{u}$ is an optimal solution of $DSN(\bar{b})$ if and only if $-\bar{u}$ is a subgradient of $v(b)$ at $b = \bar{b}$, that is $v(b) \leqq v(\bar{b}) - \bar{u}(b - \bar{b})$.

(d) every optimal solution $\bar{u}$ of $DSN(\bar{b})$ characterizes the set of all optimal solutions (if any) of $PSN(\bar{b})$ as the maximizers of

$f(x) - \bar{u}(g(x) - \bar{b})$ over K which also

satisfy the feasibility conditions $g(x) \leq \bar{b}$

and the complementary slackness conditions

$\bar{u}(g(x) - \bar{b}) = 0.$ ■

<u>Theorem 4-9</u>: If $PSN(\bar{b})$ is stable, $(\bar{x},\bar{u})$ solve $PSN(\bar{b})$ and $DSN(\bar{b})$ if and only if $(\bar{x},\bar{u})$ satisfy the following saddle-point problem:

Find $(\hat{x},\hat{u})$ such that

$$f(x)-\hat{u}(g(x)-\bar{b}) \leq f(\hat{x})-\hat{u}(g(\hat{x})-\bar{b}) \leq f(\hat{x})-u(g(\hat{x})-\bar{b})$$

for $x \in K$, $u \geqq 0$.

We call $(\hat{x},\hat{u})$ a saddle-point of $f(x) - u(g(x)-\bar{b})$. ■

<u>4-2: Nonlinear Inverse Pairs</u>

We now turn our attention to the pair of problems PVN(b) and IVN(a). For this pair we define a symmetric Lagrange function and use it to characterize optimal inverse triples. First, however, we must adjust the definitions of a value function and stability to handle our vector-valued problems.

<u>Definition 4-10</u>: The <u>value function</u> $v(b,\mu)$ associated with $PVN(b,\mu)$ for all $b \in B$ and $\mu > 0$ is defined as

$$v(b,\mu) = \sup\{\mu f(x) \mid g(x) \leqq b,\ x \in K\}.$$

The <u>value function</u> $V(a,\lambda)$ associated with $IVN(a,\lambda)$ for all $a \in A$ and $\lambda > 0$ is defined as

$$V(a,\lambda) = \inf\{\lambda g(x) \mid f(x) \geqq a,\ x \in K\}.$$ ■

Definition 4-11: $PVN(b,\mu)$ is stable at $\bar{b} \in B$ for $\mu > 0$ if $v(\bar{b},\mu)$ is finite and there exists an $M > 0$ such that

$$\frac{v(b,\mu) - v(\bar{b},\mu)}{\| b - \bar{b} \|} \leq M \text{ for all } b \in B - \{\bar{b}\}.$$

$PVN(b)$ is stable at $\bar{b} \in B$ if and only if $PVN(b,\mu)$ is stable at $\bar{b}$ for all $\mu > 0$. $IVN(a,\lambda)$ is stable at $\bar{a} \in A$ for $\lambda > 0$ if $V(\bar{a},\lambda)$ is finite and there exists an $M > 0$ such that

$$\frac{V(\bar{a},\lambda) - V(a,\lambda)}{\| a - \bar{a} \|} \leq M \text{ for all } a \in A - \{\bar{a}\}.$$

$IVN(a)$ is stable at $\bar{a} \in A$ if and only if $IVN(a,\lambda)$ is stable at $\bar{a}$ for all $\lambda > 0$. ∎

Lemma 4-7 can be extended in an obvious fashion to provide an easier method of determining the stability of $PVN(b,\mu)$ or $IVN(a,\lambda)$ for $b \in B$, $a \in A$, $\mu > 0$ and $\lambda > 0$.

For the rest of this chapter we assume (1) $PVN(b)$ is stable at all $b \in B$ and $IVN(a)$ is stable at all $a \in A$, (2) all efficient solutions of $PVN(b)$ and $IVN(a)$ are proper. Assumption (1) gives us solutions for the duals of the scalar-valued problems $PVN(b,\mu)$ and $IVN(a,\lambda)$ for all $b \in B$, $a \in A$, $\mu > 0$, $\lambda > 0$. Assumption (2) lets us find all efficient solutions to $PVN(b)$ and $IVN(a)$ by considering $PVN(b,\mu)$ and $IVN(a,\lambda)$ for $\mu > 0$, $\lambda > 0$.

Example 4-12: Let $m = 2$, $n = 2$, $k = 1$,

$f(x_1,x_2) = \sqrt{\min(x_1,x_2)}$, $g_1(x_1,x_2) = x_1$, $g_2(x_1,x_2) = x_2$,

and k be the unit rectangle $\{(x_1,x_2) \mid 1 \leq x_1 \leq 2, 1 \leq x_2 \leq 2\}$
Then PVN(b) (or in this case PSN(b)) is

$$\bar{a} = \max\{\sqrt{\min(x_1,x_2)}\}$$
$$\text{subject to } x_1 \leq b_1$$
$$x_2 \leq b_2$$
$$(x_1,x_2) \in K.$$

and IVN(a) is

$$\text{find } \bar{b} = \text{Pmin} \begin{Bmatrix} b_1 = x_1 \\ b_2 = x_2 \end{Bmatrix}$$
$$\text{such that } \sqrt{\min(x_1,x_2)} \geq a, \ (x_1,x_2) \in K.$$

For PVN(b), $v(b) = \sqrt{\min(b_1,b_2)}$. We can, without loss of generality assume that $b_1 < b_2$ and so we have that $v(b) = \sqrt{b_1}$. To check the stability of PVN(b) for any $b \in \mathcal{B}$ we must, therefore, consider the ratio $\dfrac{\sqrt{b_1+\alpha} - \sqrt{b_1}}{\alpha}$ for $\alpha > 0$. This ratio only leads to difficulties if $\alpha \to 0$ in which case we have

$\lim\limits_{\alpha \to 0} \dfrac{\sqrt{b_1+\alpha} - \sqrt{b_1}}{\alpha} = \dfrac{1}{2\sqrt{b_1}}$. However $\mathcal{B} = \{(b_1,b_2) \mid b_1 \geq 1, b_2 \geq 1\}$ so this limit is bounded regardless of what b_1 is chosen. Hence PVN(b) is stable for all $b \in \mathcal{B}$.

For IVN(a), the only efficient value is $\bar{b}_1 = \bar{b}_2 = a^2$ and for any normalized λ the optimal value of IVN(a,λ) is also a^2. Hence $V(a,\lambda) = a^2$, and all efficient points are proper. To check the stability of IVN(a) we must check the stability of IVN(a,λ) for all $\lambda > 0$. It suffices, however to check the latter only for normalized $\lambda > 0$ and for these we must consider the ratio

$$\frac{a^2 - (a-\alpha)^2}{\alpha} = +2a-\alpha \qquad \alpha>0.$$

Now we have $A = \{a|a\leq\sqrt{2}\}$, so $2a - \alpha < 2\sqrt{2}$ if $\alpha > 0$, and IVN(a,λ) is stable for all normalized $\lambda > 0$ and all $a \in A$. Thus IVN(a) is stable for all $a \in A$.■

If $\bar{a}$ is an efficient value of PVN$(\bar{b})$, $\mu\bar{a}$ is the value of PVN$(\bar{b},\mu)$ for some $\mu > 0$ at some $x \in K$. This PVN$(\bar{b},\mu)$ is stable and hence has an optimal dual solution $\bar{u}$. Hence $(\bar{x},\bar{u})$ is a saddle-point of $\mu f(x) - u(g(x)-\bar{b})$ for $x \in K$, $u \geqq 0$. Similarly if $\bar{b}$ is an efficient value of IVN$(\bar{a})$, we get an $(\bar{x},\bar{v})$ that is a saddle-point of $-\lambda g(x) + v(f(x)-\bar{a})$ for some $\lambda > 0$ and all $x \in K$, $v \geqq 0$. In the former case, since μ is constant, $\mu\bar{a}$ is constant and $(\bar{x},\bar{u})$ is a saddle-point of $\mu(f(x)-\bar{a}) - u(g(x)-\bar{b})$ for all $x \in K$, $u \geqq 0$. Similarly, in the latter case, $(\bar{x},\bar{v})$ is a saddle-point of $-\lambda(g(x)-\bar{b}) + v(f(x)-\bar{a})$ for all $x \in K$, $v \geqq 0$. This motivates the following definition.

<u>Definition 4-13</u>: Let $L(x,a,b,u,v)$ be defined as follows for the a,b,f,g of an inverse pair, PVN(b) and IVN(a):

$$L(x,a,b,u,v) = v(f(x)-a) - u(g(x)-b).$$

We call L the symmetric Lagrange function for PVN(b) and IVN(a). ■

With this definition we can obtain the following result.

<u>Theorem 4-14</u>: $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple for the inverse pair PVN(b) and IVN(a) if and only if there exist $\bar{u} > 0$, $\bar{v} > 0$ such that:

$$(1)\quad \begin{array}{ll} \text{(a)} & L(x,\bar{a},\bar{b},\bar{u},\bar{v}) \leqq L(\bar{x},\bar{a},\bar{b},\bar{u},\bar{v}) \leqq L(\bar{x},\bar{a},\bar{b},u,\bar{v}) \\ \text{(b)} & L(x,\bar{a},\bar{b},\bar{u},\bar{v}) \leqq L(\bar{x},\bar{a},\bar{b},\bar{u},\bar{v}) \leqq L(\bar{x},\bar{a},\bar{b},\bar{u},v) \end{array}$$

for all $x \in K$, $u \geqq 0$, $v \geqq 0$.

<u>Proof</u>:

■ Assume (1) holds. If we add the constant $\bar{v}\bar{a}$ to each part of (Ia), we get

$\bar{v}f(x) - \bar{u}(g(x)-\bar{b}) \leqq \bar{v}f(\bar{x}) - \bar{u}(g(\bar{x})-\bar{b}) \leqq \bar{v}f(\bar{x}) - u(g(\bar{x})-\bar{b})$

for all $x \in K$, $u \geqq 0$. Therefore $(\bar{x},\bar{u})$ is a saddle-point of $\bar{v}f(x) - u(g(x)-\bar{b})$, so $\bar{x}$ solves PVN$(\bar{b},\bar{v})$. Since $\bar{v} > 0$, $\bar{x}$ is a proper efficient solution of PVN$(\bar{b})$. Also since $\bar{u} > 0$, complementary slackness gives us $g(\bar{x}) = \bar{b}$. Similarly from (1b) we find that $f(\bar{x}) = \bar{a}$ and $\bar{x}$ is a proper efficient solution of IVN$(\bar{a})$. Therefore $\bar{a} \in A(\bar{b})$ and $\bar{b} \in B(\bar{a})$ and $(\bar{a},\bar{b},\bar{x})$ must be an optimal inverse triple.

Now assume that $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. Then there is a $\bar{\mu} > 0$ for which $\bar{\mu}\bar{a} = \bar{\mu}g(\bar{x})$ is an optimal value of PVN$(\bar{b},\bar{\mu})$ with $f(\bar{x}) = \bar{b}$. PVN$(\bar{b},\bar{\mu})$ is stable and and therefore its dual has an optimal solution $\hat{u} \geqq 0$ and and $(\bar{x},\hat{u})$ satisfies [redacted] [redacted]

$$\bar{\mu}(f(x)-\bar{a}) - \hat{u}(g(x)-\bar{b}) \leq \bar{\mu}(f(\bar{x})-\bar{a}) - \hat{u}(g(\bar{x})-\bar{b})$$
$$\leq \bar{\mu}(f(\bar{x})-\bar{a}) - u(g(\bar{x})-\bar{b})$$

for all $x \in K$, $u \geq 0$. Similarly, considering IVN$(\bar{a})$, we obtain a $\bar{\lambda} > 0$ and $\hat{v} \geq 0$ such that

$$\hat{v}(f(x)-\bar{a}) - \bar{\lambda}(g(x)-\bar{b}) \leq \hat{v}(f(\bar{x})-\bar{a}) - \bar{\lambda}(g(\bar{x})-\bar{b})$$
$$\leq v(f(\bar{x})-\bar{a}) - \bar{\lambda}(g(\bar{x})-\bar{b})$$

for all $x \in K$, $v \geq 0$. Now let $\bar{v} = \dfrac{\hat{v}+\bar{\mu}}{2}$ and $\bar{u} = \dfrac{\hat{u}+\bar{\lambda}}{2}$.

Then $\bar{v} > 0$, $\bar{u} > 0$ and $\bar{v}(f(x)-\bar{a}) - \bar{u}(g(x)-\bar{b})$ $\leq \bar{v}(f(\bar{x})-\bar{a}) - \bar{u}(g(\bar{x})-\bar{b})$. This is the left inequality of (1a) and (1b). Since $f(\bar{x}) = \bar{a}$ and $g(\bar{x}) = \bar{b}$, the right inequalities are both $0 \leq 0$. Therefore we have found a $\bar{u} > 0$ and a $\bar{v} > 0$ with the required properties. ∎

With this theorem we can prove the following result which is the nonlinear version of theorem 3-11. Therefore this result is the basis for the results of chapter 3.

<u>Theorem 4-15</u>: $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple for the inverse pair PVN$(\bar{b})$ and IVN$(\bar{a})$ if and only if there exist $\bar{\mu} > 0$ and $\bar{\lambda} > 0$ such that

(1) $\bar{x}$ solves PVN$(\bar{b},\bar{\mu})$ with optimal value $\bar{\mu}\bar{a}$,

(2) $\bar{x}$ solves IVN$(\bar{a},\bar{\lambda})$ with optimal value $\bar{\lambda}\bar{b}$

(3) $\bar{\lambda}$ solves the dual of PVN$(\bar{b},\bar{\mu})$

(4) $\bar{\mu}$ solves the dual of IVN$(\bar{a},\bar{\lambda})$

<u>Proof</u>:

∎ Assume $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. Then there exist $\bar{u} > 0$ and $\bar{v} > 0$ such that (1) in theorem

4-14 is satisfied. But, by the proof of that theorem, $(\bar{x},\bar{u})$ is a saddle-point of $\bar{v}f(x) - u(g(x)-\bar{b})$. By theorem 4-9, $(\bar{x},\bar{u})$ solves $PVN(\bar{b},\bar{v})$ and its dual. That is $\bar{x}$ solves $PVN(\bar{b},\bar{v})$ and $\bar{u}$ solves the dual of $PVN(\bar{b},\bar{v})$. Similarly $\bar{x}$ solves $IVN(\bar{a},\bar{u})$ and $\bar{v}$ solves its dual. But since $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple, $f(\bar{x}) = \bar{b}$ and $g(\bar{x}) = \bar{a}$. Therefore if we let $\bar{\mu} = \bar{v}$ and $\bar{\lambda} = \bar{u}$, we have: $\bar{x}$ solves $PVN(\bar{b},\bar{\mu})$ with optimal value $\bar{\mu}\bar{a} = \bar{\mu}f(\bar{x})$, x solves $IVN(\bar{a},\bar{\lambda})$ with optimal value $\bar{\lambda}\bar{b} = \bar{\lambda}g(\bar{x})$, $\bar{\lambda}$ solves the dual of $PVN(\bar{b},\bar{\mu})$, and $\bar{\mu}$ solves the dual of $IVN(\bar{a},\bar{\lambda})$.

Assume conditions (1) to (4) are satisfied for some $\bar{\mu} > 0$, $\bar{\lambda} > 0$. Therefore by theorem 4-9 $(\bar{x},\bar{\lambda})$ is a saddle-point of $\bar{\mu}f(x) - \lambda(g(x)-\bar{b})$ for $x \in K$, $\lambda \geqq 0$, and $(\bar{x},\bar{\mu})$ is a saddle-point of $\mu(f(x)-\bar{a}) - \bar{\lambda}g(x)$ for $x \in K$, $\mu \geqq 0$. Thus the inequalities in (1) of theorem 4-14 hold and $(\bar{a},\bar{b},\bar{x})$ must be an optimal inverse triple.∎

<u>Corollary</u>: Theorem 3-11.∎

In the nonlinear case we cannot obtain a complete equivalent of theorem 3-12. In the linear case when we started with a multiplier vector $\bar{\mu}$ and an estimated $\bar{b}$ we found an optimal inverse triple by adjusting $\bar{b}$ until we could find a positive solution to the dual of $PVL(\bar{b},\bar{\mu})$. We argued that, because of linearity, we only needed to consider a finite number of possible alternatives for $\bar{b}$. In the nonlinear case we cannot use this argument and, thus, we may not be able to develop a method of adjusting $\bar{b}$ in a finite number of steps to a value that allows a positive solution for the dual of $PVN(\bar{b},\bar{\mu})$. However we do have the following partial result.

<u>Theorem 4-16</u>: If $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple then:

(1) if $\bar{a}$ is an efficient value of $PVN(\bar{b})$ found by solving $PVN(\bar{b},\mu)$ for some $\mu > 0$, then $\bar{b}$ is an efficient value of $IVN(\bar{a})$ found by solving $IVN(\bar{a},\lambda)$ where λ is any positive optimal solution of the dual of $PVN(\bar{b},\mu)$.

(2) if $\bar{b}$ is an efficient value found by solving $IVN(\bar{a},\lambda)$ for some $\lambda > 0$, then $\bar{a}$ is an efficient value found by solving $PVN(\bar{b},\mu)$ where μ is any positive optimal solution of the dual of $IVN(\bar{a},\lambda)$.

<u>Proof</u>:

∎ (1) Since $\mu > 0$ and $\lambda > 0$, they provide the $\bar{u}$ and $\bar{v}$ needed in theorem 4-14. So, as in the proof of that theorem, $\bar{b}$ is the efficient value found by solving $IVN(\bar{a},\lambda)$. (2) Similar.∎

<u>Corollary</u>: Theorem 3-5 part (1).

<u>Proof</u>:

∎ Since $PSL(\bar{b})$ is scalar-valued, we do not need a μ. $\bar{a}$ is the efficient value found by solving $PSN(\bar{b})$ since $(\bar{a},\bar{b},\bar{x})$ is an optimal inverse triple. Therefore $\bar{b}$ is an efficient value of $IVL(\bar{a})$ generated by solving $IVL(\bar{a},\bar{y})$ since $\bar{y} > 0$ solves the dual of $PSL(\bar{b})$. ∎

This result gives us part of what we had before. If we start with $\bar{\mu}$ and $\bar{b}$ and if the dual of $PVN(\bar{b},\bar{\mu})$ has a positive solution we can obtain an optimal inverse triple. However, if not, we cannot guarantee that we can find an adjusted $\bar{b}$ for which we can obtain a positive solution. For particular types of problems, as in the linear case, algorithms may exist because the characteristics of the problems allow us to make extra assumptions. However, in general this cannot be done.

Bibliography

1. Balinski, M.L. and W.J. Baumol; "The dual in nonlinear programming and its economic interpretation", Rev Eco Studies 35 (1968) 237-256.

2. Benayoun, R., J. de Montgolfier, J. Tergny and O. Laritchev; "Linear programming with multiple objective functions: step method (STEM)", Math Prog 1 (1971) 366-375.

3. Cass, D.; "Duality: a symmetric approach from the economist's vantage point"; J Eco Th 7 (1974) 272-295.

4. Cassidy, R.G.; Ph.D. Thesis, On Random Payoff Games, Dalhousie University, Halifax, Nova Scotia, 1970.

5. Cassidy, R.G., C.A. Field and W.R.S. Sutherland; "Inverse programming: theory and examples", Math Prog 4 (1973) 297-308.

6. Charnes, A. and W.W. Cooper: Management Models and Industrial Applications of Linear Programming, Vol. 1, Wiley, New York, 1951.

7. Cochrane, J.L. and M. Zeleny, eds., Multiple Criteria Decision Making, Un. of South Carolina Press, Columbia, S.C., 1973.

8. DaCunha, N.O. and E. Polak; "Constrained minimization under vector-valued criteria in finite dimensional

space", J Math Anal Appl 19 (1967) 103-124.

9. Ecker, J.G. and I.A. Kouada; "Finding efficient points for linear multiple objective programs", Math Prog 8 (1975) 375-377.

10. Evans, J.P. and F.J. Gould; "A nonlinear duality theorem without convexity", Econometrica 40 (1972) 487.

11. Evans, J.P. and R.E. Steuer; " A revised simplex method for linear multiple objective programs", Math Prog 5 (1973) 54-72.

12. Gale, D.; "A geometric duality theorem with economic applications", Rev Eco Studies 34 (1967) 19-24.

13. Geoffrion, A.M.; "Duality in nonlinear programming: a simplified applications - oriented development", SIAM Review 13 (1971) 1-37; reprinted in A.M. Geoffrion, ed.; Perspectives on Optimization: A Collection of Expository Articles, Addison-Wesley, Reading, Mass., 1972.

14. Geoffrion, A.M.; "Proper efficiency and the theory of vector maximization", J Math Anal Appl 22 (1968) 618-630.

15. Isermann, H.; "Proper efficiency and the linear vector maximization problem", Ops Res 22 (1974) 189-191.

16. Karlin, S.; Mathematical Methods and Theory in Games, Programming and Economics, Vol.1, Addison-Wesley, Reading, Mass., 1959.

17. Klinger, A.; "Improper solutions of the vector maximization problem", Ops Res 15 (1967) 570-572.

18. Klinger, A.; "Vector-valued performance criteria", IEEE Trans Auto Control, AC-9 (1964) 117-118.

19. Koopmans, T.C. ed.; Activity Analysis of Production and Allocation, Wiley, New York, 1951.

20. Kornbluth, J.S.H.; "Multiple objective linear programming", ORQ 25 (1974) 599-614.

21. Kuhn, H.W. and A.W. Tucker; "Nonlinear programming", in J.Neyman, ed., Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Un. of California Press, Berkeley, California, 1950.

22. MacCrimmon, K.R.; "An overview of multiple objective decision making", p.18 in [7].

23. Mangasarian, O.L.; Nonlinear Programming, McGraw-Hill, New York, 1969.

24. Mangasarian, O.L. and J. Ponstein; "Minimax and duality in nonlinear programming", J Math Anal Appl 11 (1965) 504-518.

25. Philip, J.; "Algorithms for the vector maximization

problem", Math Prog 2 (1972) 207-229.

26. Rockafellar, R.T.; "Duality in nonlinear programming", in G.B. Dantzig and A.F. Veinott, eds.; Mathematics of the Decision Sciences, Part 1, Am Math Society, Providence, R.I., 1968.

27. Roy, B.; "Problems and methods with multiple objective functions", Math Prog 1 (1971) 239-266.

28. Stoer, J.; "Duality in nonlinear programming and the minimax theorem", Numer Math 5 (1963) 371-379.

29. Wolfe, P.; "A duality theorem for non-linear programming", Quart Appl Math 19 (1961) 239-244.

30. Yu, P.L.; "Cone convexity, cone extreme points and nondominated solutions in decision problems with multiobjectives", JOTA 14 (1974) 319-377.

31. Yu, P.L.; "Introduction to domination structures in multicriteria decision problems", in [7].

32. Yu, P.L. and M. Zeleny; "The set of all nondominated solutions in linear cases and a multicriteria simplex method", J Math Anal Appl 49 (1975) 430-468.